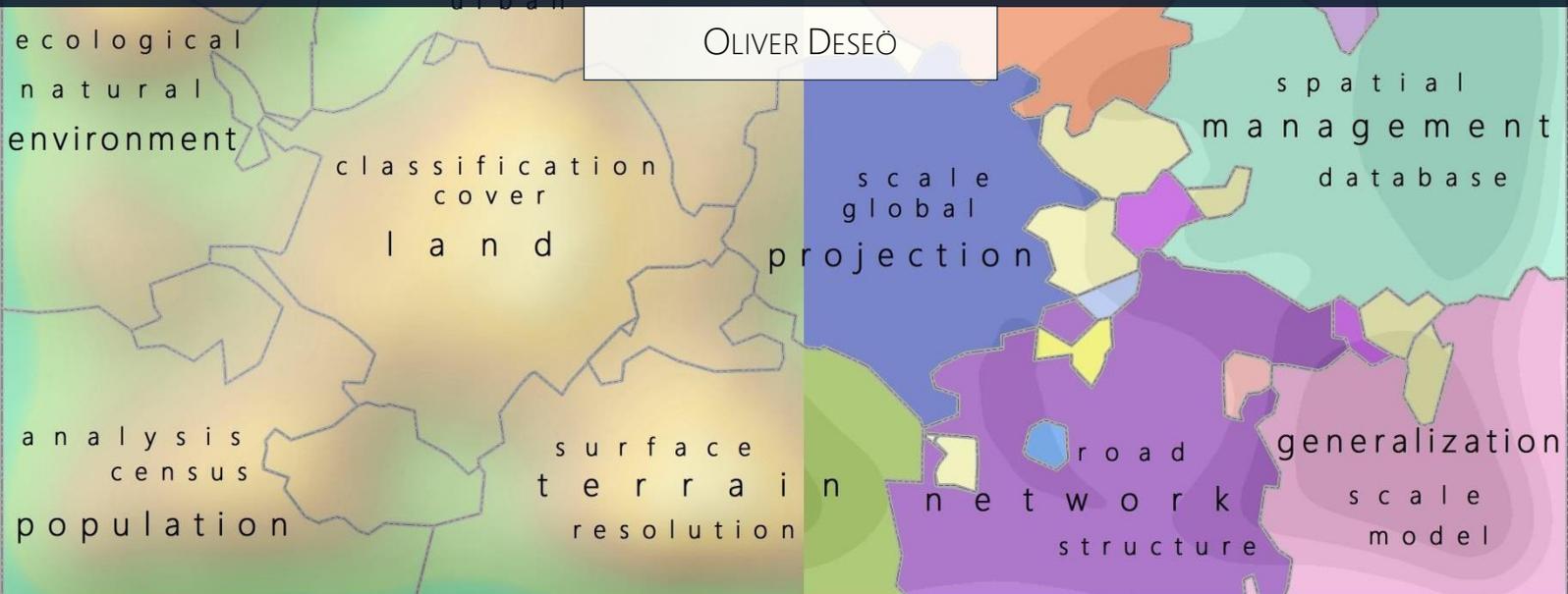


# INFORMATIONSLANDSCHAFTEN IM VERGLEICH: THEMATISCHE KARTE VS. RELIEFKARTE



2015

Universität Zürich  
Geographisches Institut



# Informationslandschaften im Vergleich: Thematische Karte Vs. Reliefkarte

---

Eine empirische Studie vergleicht geographische Metaphern bei Verräumlichungen  
basierend auf Konferenzbeiträgen.

Masterarbeit GEO 511  
Wil SG, 30.09.2015

VERFASSER:

**OLIVER DESEÖ**

OERLIKONERSTRASSE 40

CH - 8057 ZÜRICH

10-756-633

oli\_deseoe@hotmail.com

BETREUUNG UND LEITUNG:

**PROF. DR. SARA IRINA FABRIKANT**

ABTEILUNG GIVA

GEOGRAPHISCHES INSTITUT UZH

WINTERTHURERSTRASSE 190

CH – 8057 ZÜRICH

BETREUUNG:

**ANDRÉ BRUGGMANN**

ABTEILUNG GIVA

GEOGRAPHISCHES INSTITUT UZH

WINTERTHURERSTRASSE 190

CH – 8057 ZÜRICH

Geographisches Institut der Universität Zürich,  
Abteilung: Geographische Informationsvisualisierung & –Analyse (GIVA)

---



## VORWORT

---

Diese Masterarbeit repräsentiert den Abschluss meines Masterstudiums am Geographischen Institut der Universität Zürich. Gleichzeitig bündelt sie das im Verlauf des Studiums angesammelte Wissen über diverse Techniken und Theorien. Natürlich ist es schwierig, die ganze Vielfalt der Geographie in eine Arbeit zu stecken, weshalb sich meine Masterarbeit mit der gewählten Spezialisierung des Masterstudiengangs beschäftigt. Geographische Informationssysteme und vor allem auch die Visualisierungen, die mit diesen zustande kommen, haben mich in meinem fünfjährigen Studium immer fasziniert. Da man im Studium immer gewisse Vorgaben bei der Darstellung von räumlichen Daten hat, interessierte mich die Idee, andere – also keine räumlich referenzierbaren – Informationen nach ähnlichen Prinzipien wie bei gewöhnlichen, alltäglichen Karten darzustellen. Dabei inspirierten mich andere Arbeiten wie eine Landschaft der Musikrichtungen (Biberstine et al. 2010) aber auch der Wissensdrang, einen solchen Visualisierungsprozess eigenständig durchzuführen. Aus dieser Idee entstand ein Vergleich von Informationslandschaften, welche in ihrer Darstellung an thematische Karten oder Reliefkarten angelehnt sind. Neben dem Darstellungsprozess, welcher von der Datenaufbereitung über verschiedene Methoden und Techniken bis zum finalen Produkt einer solchen Informationslandschaft durchgezogen wurde, soll eine Studie diese Produkte vergleichen und analysieren.

Während dem ganzen Visualisierungsprozess, aber auch während der Studienvorbereitung und deren Auswertung, haben verschiedene Personen – Freunde, Familie und Verwandte - zum Gelingen dieser Arbeit beigetragen und mich immer unterstützt. Ganz besonders danken möchte ich folgenden Personen.

- › Prof. Dr. Fabrikant für diverse methodische und inhaltliche Tipps und Ideen sowie für die Kommunikation mit der Internationalen Kartographischen Vereinigung
- › André Bruggmann für die Unterstützung bei allen wichtigen Fragen im methodischen und inhaltlichen Vorgehen und für die diversen Rückmeldungen und Besprechungen zur Arbeit
- › Prof. Dr. André Skupin für die methodische Unterstützung beim Visualisierungsprozess
- › Patrice Frei und Ramón Huldi für das beinahe tägliche Auseinandersetzen mit Fragen aller möglichen Art und die entspannte Atmosphäre im Y25-H86
- › Dem engsten Familien- und Freundeskreis für das Korrekturlesen, die zahlreichen motivierenden Gespräche und verschiedenste Arten der Unterstützung
- › Allen Umfrageteilnehmenden für deren Teilnahme, für die sorgfältige Bearbeitung des Fragebogens und für die vielen Rückmeldungen zur Studie

Wil SG, 27.09.2015

Oliver Deseö

---



## ABSTRACT

---

Die Unübersichtlichkeit der tagtäglich wachsenden Datenmenge strebt nach neuen Visualisierungsmöglichkeiten, welche das Extrahieren der relevanten Dateneigenschaften vereinfacht. Mit dem wissenschaftlichen und technologischen Fortschritt sind die Möglichkeiten der Informationsvisualisierung in neue Dimensionen vorgestossen. Ein häufig angewandter Grundsatz zeigt sich im ersten Gesetz der Geographie von Tobler (1970), welches die Ähnlichkeit von Dingen im Raum abhängig von der räumlichen Distanz dazwischen macht. Basierend auf diesem Prinzip kann man die Übersicht von nicht-räumlichen Daten analog zu räumlich referenzierten Informationen neu strukturieren und ordnen. Diese Methode nennt sich Verräumlichung und ist das Fundament dieser Arbeit. Wenn man nun diese nicht-räumlichen Daten in einen Informationsraum projiziert, als ob den Daten eine räumliche Verteilung zugrunde liegt, dann kann man dieselben Darstellungsprinzipien wie bei gewöhnlichen kartographischen Produkten zur entsprechenden Visualisierung nutzen. Diese Arbeit beschreibt einerseits einen Visualisierungsprozess solcher Informationslandschaften und zeigt zwei Lösungen, die sich an einer thematischen Karte oder einer Reliefkarte am Beispiel des Schweizer Weltatlas (IKG ETH Zürich 2015) orientieren. Dabei stellt die thematisch klassierte Informationslandschaft dieselben Informationen wie die kontinuierlich dargestellte Reliefkarte der Grundlagedaten. Diese bestehen aus Konferenzbeiträgen der biennalen Konferenzen der Internationalen Kartographischen Vereinigung von 1993 bis 2013, welche über die Methoden des Topic Modeling und der Selbstorganisierende Karten (*SOM*) in den Informationsraum projiziert werden. Die beiden resultierenden Informationslandschaften wurden aufgrund verschiedener geographischer Metaphern in einer Studie untersucht. Dabei spielen sowohl die Ähnlichkeitswahrnehmung als auch die Übersicht eine entscheidende Rolle. Die grundlegende Frage war dabei, ob das Übertragen der geographischen Metaphern wie die Region oder die Distanz für beide Informationslandschaften funktioniert und welche davon den grössten Einfluss auf die Wahrnehmung hat. Die empirische Evaluation ergab keinen signifikanten Unterschied nur zwischen den beiden Visualisierungstypen. Nur die Oberflächenbeschaffenheit wird in der Informationslandschaft, die auf einer Reliefkarte basiert, als bedeutend einflussreicher bewertet. Mit dem Einbezug des geographischen Hintergrunds der Testpersonen kann man dabei einzelne signifikante Differenzen zu erkennen. Die Distanz hat bei den Geographinnen und Geographen einen grösseren Einfluss als bei den Nicht-Geographinnen und Nicht-Geographen. Damit auch Laien solche Darstellungen problemlos verstehen sind einzelne Anpassungen in der Informationslandschaft oder eine klare Legende nötig. Dennoch erweist sich eine solche Informationslandschaft als gestandene Darstellungsmethode von nicht-räumlichen Daten.

### Schlüsselwörter

SOM, Topic Modeling, Informationslandschaft, Informationsvisualisierung, Ähnlichkeitswahrnehmung

---



# INHALT

---

<b>1</b>	<b>Einleitung</b> .....	<b>1</b>
<b>2</b>	<b>Theorie und Forschungsstand</b> .....	<b>3</b>
2.1	Der nicht-geographische Informationsraum.....	3
2.1.1	Grundlage: Textdaten.....	3
2.1.2	Die Verräumlichung.....	5
2.1.3	Self-Organizing-Maps (SOM) .....	8
2.1.4	Diskret vs. Kontinuierlich.....	9
2.2	Visualisierungsbeispiele und veröffentlichte Studien .....	10
2.2.1	Visualisierungsbeispiele.....	10
2.2.2	Veröffentlichte Studien.....	13
<b>3</b>	<b>Forschungsfragen und Methodik</b> .....	<b>15</b>
3.1	Arbeitsziele .....	15
3.2	Fragestellungen & Hypothesen.....	16
3.2.1	Fragestellungen.....	16
3.2.2	Hypothesen .....	16
3.3	Daten: Grundlage & Aufbereitung .....	17
3.3.1	Verarbeitung der Daten (Topic Modeling) .....	18
3.4	Visualisierungsprozess.....	21
3.4.1	Allgemeiner Arbeitsprozess.....	22
3.4.1.1	Self-Organizing Maps (SOM).....	22
3.4.1.2	Regionen in der Informationslandschaft.....	24
3.4.1.3	Beschriftung der Regionen in der Informationslandschaft .....	27
3.4.1.4	Topographie der Informationslandschaft.....	29
3.4.2	Visualisierungstyp D.....	31
3.4.3	Visualisierungstyp K.....	32
3.4.4	Zeitliche Visualisierung.....	33
3.5	Studiendesign .....	35
3.5.1	Stichprobe .....	35
3.5.2	Web-Plattform.....	35
3.5.3	Studienaufbau.....	35
3.5.3.1	Pretest .....	35
3.5.3.2	Hauptexperiment.....	36
3.5.3.3	Posttest.....	37
3.5.4	Experimentelles Design.....	38
3.5.4.1	Versuchsdurchführung.....	38
3.5.4.2	Unabhängige Variable.....	39
3.5.4.3	Abhängige Variablen.....	40
3.5.4.3.1	Leistung.....	40
3.5.4.3.2	Sicherheit.....	41
3.5.4.3.3	Präferenz.....	41
3.5.4.4	Kontrollvariable.....	41
3.5.4.4.1	Geographie-Studium.....	41
3.5.5	Statistik .....	42

---

<b>4</b>	<b>Resultate</b> .....	43
4.1	Stichprobe.....	43
4.2	Visualisierungstyp.....	45
4.2.1	Übersicht.....	45
4.2.2	Ähnlichkeit.....	46
4.2.3	Aufgabenübersicht (Ähnlichkeit).....	48
4.2.4	Textantworten.....	50
4.2.5	Präferenz.....	51
4.3	Geographische Ausbildung.....	52
4.3.1	Übersicht.....	52
4.3.2	Ähnlichkeit.....	53
4.3.3	Präferenz.....	55
4.4	Zusammenfassung der Resultate.....	56
<b>5</b>	<b>Diskussion</b> .....	57
5.1	Leistung.....	57
5.1.1	Performance.....	57
5.1.1.1	Übersichtsaufgaben.....	57
5.1.1.2	Ähnlichkeit.....	58
5.1.2	Rückschlüsse zur ersten Forschungsfrage.....	59
5.1.3	Sicherheit.....	59
5.1.4	Rückschlüsse zur zweiten Forschungsfrage.....	59
5.2	Präferenz.....	60
5.2.1	Rückschlüsse zur dritten Forschungsfrage.....	60
5.3	Hintergrundwissen.....	61
5.3.1	Übersichtsaufgaben.....	61
5.3.2	Ähnlichkeit.....	61
5.3.3	Präferenz.....	61
5.3.4	Rückschlüsse zur vierten Forschungsfrage.....	61
5.4	Verbesserungsmöglichkeiten für Informationslandschaften.....	62
5.5	Limitierungen und Probleme der Studie.....	63
<b>6</b>	<b>Zusammenfassung &amp; Ausblick</b> .....	65
<b>7</b>	<b>Literatur &amp; Quellen</b> .....	67
<b>Anhang</b>	.....	73

---

## ABBILDUNGEN

---

Abbildung 1-1: Ausschnitt der europäischen Landschaft und Ausschnitt der Musiklandschaft.....	2
Abbildung 2-1: Geographische Informationslandschaften basierend auf Textdaten.....	3
Abbildung 2-2: Verfeinertes Prinzip der Verräumlichung .....	7
Abbildung 2-3: Verschiedene Darstellungen einer Musiklandschaft.....	10
Abbildung 2-4: Geographische Landschaft basierend auf AAG-Konferenzbeiträge.....	11
Abbildung 2-5: ICA-Landschaft der Konferenzen in Peking (2001) und Durban (2003).....	11
Abbildung 2-6: Verkleinerte Ansicht der kompletten Medizinlandschaft .....	12
Abbildung 3-1: Arbeitsziele.....	15
Abbildung 3-2: Schema des Topic Modeling.....	18
Abbildung 3-3: Relative Häufigkeit der einzelnen Topics in der gesamten Datengrundlage .....	20
Abbildung 3-4: Durchschnittliche Wahrscheinlichkeit pro Topic.....	20
Abbildung 3-5: Flussdiagramm der verschiedenen Visualisierungsprozesse.....	21
Abbildung 3-6: SOM-Output .....	23
Abbildung 3-7: Resultat K-Means Clusteranalyse .....	25
Abbildung 3-8: Quantitative Generalisierung.....	27
Abbildung 3-9: Arbeitsprozess von Punktdaten zu Regionen .....	26
Abbildung 3-10: Beschriftung der Regionen.....	29
Abbildung 3-11: Begriffsdominanz-Landschaft.....	30
Abbildung 3-12: Ausschnitt aus einer politisch-gegliederten, europäischen Karte.....	31
Abbildung 3-13: Finale diskretisierte Visualisierung.....	31
Abbildung 3-14: Ausschnitt aus einer europäischen Reliefkarte.....	32
Abbildung 3-15: Finale kontinuierliche Visualisierung.....	32
Abbildung 3-16: Schwerpunktverschiebung der ICA-Konferenzbeiträge von 1993-2013 .....	33
Abbildung 3-17: Verteilung der Konferenzbeiträge nach Jahr.....	34
Abbildung 3-18: Hauptexperiment.....	36
Abbildung 3-19: Beispielsvisualisierungen für Fragentyp «Übersicht» .....	36
Abbildung 3-20: Beispielsvisualisierungen für Frage «Ähnlichkeit».....	37
Abbildung 3-21: Boxplot zur durchschnittlichen Erfahrung abhängig von der Ausbildung.....	41
Abbildung 4-1: Aufteilung der Klasse «Andere» in der Kategorie «Ausbildung».....	43
Abbildung 4-2: Aufteilung der Stichprobe .....	43
Abbildung 4-3: Aufteilung der Stichprobe nach eingeschätzten Kenntnissen.....	44
Abbildung 4-4: Aufteilung der geographischen Spezialisierungen des Hauptfachstudiums .....	44
Abbildung 4-5: Vergleich der korrekten Antworten für die Übersichtsaufgaben.....	45
Abbildung 4-6: Vergleich der Sicherheit beim Beantworten der Übersichtsaufgaben.....	45
Abbildung 4-7: Vergleich der Ähnlichkeit beim Beantworten der Ähnlichkeitssaufgaben.....	46
Abbildung 4-8: Vergleich des Einfluss der einzelnen Faktoren .....	47
Abbildung 4-9: Vergleich der Sicherheit beim Beantworten der Ähnlichkeitsaufgaben.....	47
Abbildung 4-10: Vergleich der Kommentare der Testpersonen.....	50
Abbildung 4-11: Vergleich der Präferenz der Testpersonen .....	51
Abbildung 4-12: Vergleich der Ausbildung bei Übersichtsaufgaben.....	52

---

Abbildung 4-13: Vergleich der Ausbildung für die Sicherheit bei Übersichtsaufgaben.....	52
Abbildung 4-14: Vergleich der Ausbildung für die Ähnlichkeitswahrnehmung.....	53
Abbildung 4-15: Vergleich der Ausbildung auf den Einfluss der einzelnen Faktoren.....	54
Abbildung 4-16: Vergleich der Ausbildung für die Sicherheit bei Ähnlichkeitsaufgaben.....	54
Abbildung 4-17: Vergleich der Ausbildung bei der Präferenz der Testpersonen.....	55

## TABELLEN

---

Tabelle 2-1: Grobes Prinzip der Verräumlichung.....	5
Tabelle 3-1: Forschungsfragen.....	16
Tabelle 3-2: Hypothesen.....	16
Tabelle 3-3: Semi-strukturierte Grundlagedaten.....	17
Tabelle 3-4: Vereinfachte exemplarische Artikel-Themen-Matrix I.....	19
Tabelle 3-5: Vereinfachte Themenübersicht aus Topic Modeling.....	19
Tabelle 3-6: Vereinfachte exemplarische Artikel-Themen-Matrix II.....	24
Tabelle 3-7: Artikel-Ähnlichkeits-Matrix.....	24
Tabelle 3-8: Ausschnitt aus der Vorstudie.....	28
Tabelle 3-9: Reihenfolge der Visualisierungstypen für beide Fragebogen.....	38
Tabelle 3-10: Experimentelles Design.....	40
Tabelle 3-11: Skalenübersicht.....	42
Tabelle 4-1: Statistische Auswertung der einzelnen Faktoren.....	47
Tabelle 4-2: Übersicht zur Beantwortung der einzelnen Fragen.....	49
Tabelle 4-3: Vergleich der einzelnen Faktoren für Aufgabe 5.....	49
Tabelle 4-4: Matrix der korrekten Antworten der Ähnlichkeitsfragen.....	53
Tabelle 4-5: Statistische Auswertung der einzelnen Faktoren.....	54

## ABKÜRZUNGEN

---

<i>AAG</i>	Association of American Geographers (deutsch: Vereinigung der Amerikanischen Geographinnen und Geographen)
<i>BMU</i>	Best-Matching Unit (deutsch: Am besten passendste Einheit)
<i>GIS</i>	Geographische Informationssysteme
<i>ICA</i>	International Cartographic Association (deutsch: Internationale Kartographische Vereinigung)
<i>MDS</i>	Multi-Dimensionale Skalierung
<i>NLTK</i>	National Language Toolkit (Bird et al. 2009)
<i>NWB</i>	Network Workbench (Software)
<i>PCA</i>	Principal Component Analysis (deutsch: Hauptkomponentenanalyse)
<i>SOM</i>	Self-Organizing Map (deutsch: Selbst-Organisierende Karte)

---

## 1 EINLEITUNG

---

Wegen der endlos wachsenden Datenmengen sind alternative Visualisierungen auch von nicht-geographischen Informationen interessanter denn je (Buckley et al. 2000). Durch die stetig wachsenden Datenberge wird die Extraktion von relevanten Informationen immer komplexer (Fabrikant & Buttenfield 2001; Keim et al. 2005). Als mögliche Lösung der Erforschung solch grosser Datensätze können Informationsvisualisierungstechniken dienen (Keim et al. 2005), welche mit dem technologischen Fortschritt zunehmend Möglichkeiten in der verständlichen Darstellung von Informationseigenschaften bieten (Skupin & Hagelman 2005). Dabei sollen auch die traditionelle Kartographie und Transformationstechniken, die aus der geographischen Informationswissenschaft abgeleitet werden, auf diverse Aspekte der Informationsvisualisierung angewandt werden (Skupin 2004). Die Informationsvisualisierung beschäftigt sich mit der Kunst und Technologie der Gestaltung und Implementierung von interaktiven, computerbasierenden Werkzeugen zur Wissensgewinnung in grossen, nicht-räumlichen Datenbanken. Die graphische Verräumlichung als Visualisierungsmethode von nicht-räumlichen Informationen unterscheidet sich von der gewöhnlichen Datenvisualisierung und Geovisualisierung dadurch, dass die Interpretation der Daten analog zu räumlichen Daten durchgeführt werden kann (Fabrikant & Skupin 2005). So kann man in einer solchen Verräumlichung enthaltene Informationen in grossen relationalen Datenbanken, wie man sie im Internet findet, auf ihre Beziehungen zueinander untersuchen (Fabrikant & Salvini 2011). Dabei wendet man eine räumliche Metapher an, welche nicht-räumliche Informationen in einem räumlichen oder geographischen Gebiet darstellt (Fabrikant et al. 2004). Die Eingebundenheit, welche durch das interaktive kognitive Eintauchen des Betrachters in diese virtuelle Umgebung benötigt wird, und die Verortung, welche Eigenschaften der geographischen Welt in diese fiktive Landschaft übertragen kann, sind Voraussetzungen einer solchen Visualisierung (Couclelis 1998).

Ausser der räumlichen Metapher benötigt eine virtuelle Informationslandschaft weitere geographische Prinzipien und bei der finalen Darstellung kartographische Techniken (Skupin & Fabrikant 2003). Das wichtigste geographische Prinzip findet sich im Ersten Gesetz der Geographie von Tobler (1970) wieder. In diesem wird die Abhängigkeit der Verwandtschaft einzelner Dinge von der räumlichen Distanz beschrieben. Befinden sich Informationen im Raum näher beieinander, sind sie sich generell ähnlicher. Dieses Gesetz lässt sich auf die nicht-geographische Informationslandschaft übertragen. Zu den kartographischen Techniken gehören die Prozesse der Generalisierung und das entsprechende Anwenden von visuellen Variablen zur Darstellung einzelner Repräsentationen (Bertin 1967). Diese gelten als Sätze der graphischen Sprache und unterscheiden sich in ihrer Aussagekraft und Bedeutung (Mackinlay 1986).

Um nun eine enorme Datenmenge zu verarbeiten und die darin enthaltenen Informationen und Beziehungen übersichtlich und verständlich zu visualisieren, müssen zunächst die Dimensionen der Daten so reduziert werden, dass die wichtigsten Eigenschaften der Datenmenge vorhanden bleiben und nebensächliche Datenrelationen für eine verbesserte Übersicht vernachlässigt werden können (Skupin & Fabrikant 2003). Neben den verschiedenen Möglichkeiten bei der Reduktion der Datendimensionen gibt es unterschiedliche Darstellungsvarianten für einzelne Eigenschaften der in der Datenmenge enthaltenen Informationen. Ein immer wiederkehrender Forschungsgegenstand in der Geographie ist dabei der Unterschied zwischen Raster- oder Vektordatendarstellungen – also das Verständnis von kontinuierlichen Oberflächen wie zum Beispiel die Topographie oder einzelne diskrete geometrische Objekte, die man als Punkt, Linien oder Polygone visualisieren kann (Couclelis 1992).

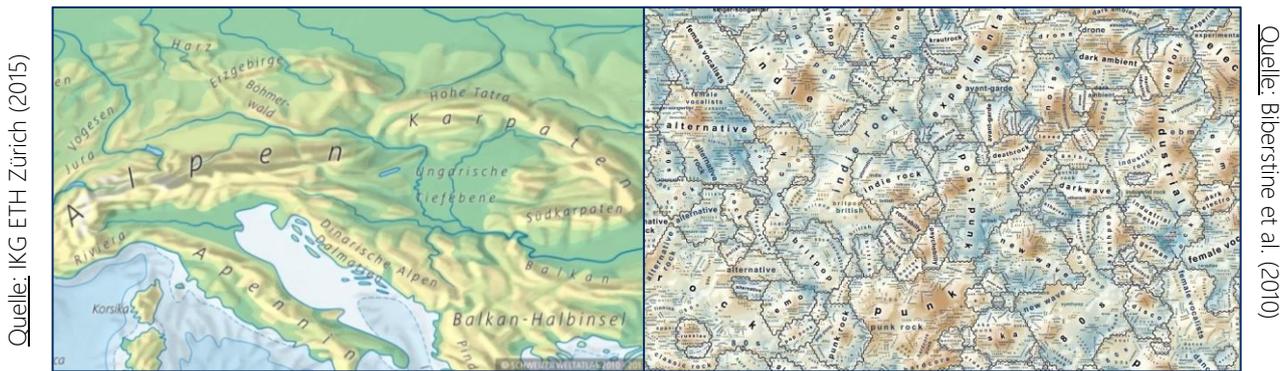


Abbildung 1-1: Ausschnitt der europäischen Landschaft und Ausschnitt der Musiklandschaft

Als Beispiel erkennt man in Abbildung 1-1(links) einen Ausschnitt aus dem Schweizer Weltatlas<sup>1</sup>, der mittels diskreten und kontinuierlichen Elementen einen Teil der europäischen Landschaft darstellt. In der musikalischen Informationslandschaft (rechts in Abbildung 1-1), welche aus user-generierten Tags der Musikplattform Last.fm<sup>2</sup> extrahiert wurde und so einen Überblick über die Musiklandschaft mit all ihren verschiedenen Richtungen und Stilen verspricht, kombiniert man diskrete und kontinuierliche Elemente. Unklar dabei ist, ob man Darstellungsmethoden von georeferenzierten Landschaften auf nicht-räumliche Informationsvisualisierungen übertragen kann und inwiefern die kontinuierlichen und/oder diskreten Elemente die Wahrnehmung einer solchen Informationslandschaft beeinflussen. Diese unterschiedlichen Visualisierungen sollen in dieser Arbeit der modernen Literatur der Informationsvisualisierung und -analyse entsprechend gestaltet und durch kartographische Methoden optimiert werden. Daraus ergeben sich folgende offene Fragestellungen:

- Kann man eine Informationslandschaft wie eine politisch gegliederte oder wie eine Relief-Karte darstellen? Wenn ja, gibt es Interpretationsunterschiede?
- Welche geographischen Metaphern beeinflussen die Interpretation der Informationslandschaft?
- Gibt es sonstige Faktoren, welche die Interpretation der Informationslandschaft beeinflussen?

Um diesen beiden Fragestellungen nachzugehen, müssen zuerst die Visualisierungen über Datenbankverarbeitungsprogramme auch für das Kreieren solcher Darstellungen entwickelte Visualisierungsprogramme und -Tools erstellt werden. Die kreierte Visualisierungen werden in einer Studie verglichen.

Der Hauptteil der Arbeit ist in fünf Abschnitte gegliedert. Nach der kurzen Einführung in die gewählte Thematik wird im Kapitel 2 in drei Unterkapiteln die theoretische Grundlage der gewählten Methodik mit dem aktuellen Forschungsstand erklärt und verschiedene Visualisierungsbeispiele aufgezeigt. Anschliessend wird die methodische Anleitung im Kapitel 0, die zur Beantwortung der im selben Kapitel beschriebenen Forschungsfragen führen sollen. Die Auswertung der empirisch gewonnenen Daten, welche die zu den Forschungsfragen formulierten Hypothesen untersuchen, werden im Kapitel 4 dargelegt. Diese Ergebnisse werden im Kapitel 5 sowohl im Zusammenhang mit den Forschungsfragen interpretiert, mit bereits getätigten Arbeiten verglichen als auch ihre Relevanz für die Forschung diskutiert. Die Schlussfolgerung fasst im Kapitel 0 die Ergebnisse nochmals zusammen und gewährt einen Ausblick in zukünftige Forschung.

<sup>1</sup> <http://schweizerweltatlas.ch/> (Letzter Zugriff: 15.09.2015)

<sup>2</sup> <http://www.last.fm/> (Letzter Zugriff: 12.09.2015)

## 2 THEORIE UND FORSCHUNGSSTAND

Wie schon im Kapitel «Einleitung» beschrieben wurde beschäftige ich mich mit unterschiedlichen Informationsvisualisierungsarten. Im folgenden Kapitel werde ich einerseits die verwendete Datengrundlage beschreiben und die verwendete Visualisierungsmethode im nicht-geographischen Informationsraum erklären. Dabei stehen Unterschiede zwischen diskreten und kontinuierlichen Daten und Darstellungen im Mittelpunkt, da dies auch ein signifikanter Unterschied zwischen politisch-gegliederten und Relief-Karten ist. Andererseits werden die in dieser Arbeit verwendeten Visualisierungstypen mit bereits vorhandenem Hintergrundwissen erklärt und begründet. Ausserdem wurden in diesem Gebiet schon einige Studien veröffentlicht, die auch in dieser Arbeit bedeutend sind.

### 2.1 DER NICHT-GEOGRAPHISCHE INFORMATIONSRAUM

Damit die nicht-räumliche Textgrundlage überhaupt visualisiert werden kann, ist ein unterstützendes Konstrukt nötig. Der Informationsraum verfügt über klar definierte Strategie im Organisieren solcher Informationen und kann mit logischen, metrischen, nicht-metrischen oder einprägsamen Koordinaten formalisiert werden (Skupin & Buttenfield 1997). Er nutzt dabei verschiedene (geographische) Metaphern, um auch nicht-räumliche Daten wie zum Beispiel Konferenzbeiträge darstellen zu können. Doch bevor man die Eingangsdaten in den Informationsraum projizieren kann muss man diese zuerst sinnvoll aufbereiten. Dieser Abschnitt verweist auf Forschungsarbeiten, welche von der Textdatenaufbereitung bis zu deren Projektion in den nicht-geographischen Informationsraum von Bedeutung sind. Dabei spielt auch die Verräumlichung eine zentrale Rolle.

#### 2.1.1 GRUNDLAGE: TEXTDATEN

Für die Gestaltung einer solchen Informationslandschaft gibt es verschieden mögliche Datengrundlagen. So kann man beispielsweise digital verfügbare Texte statistisch auswerten und so Informationen über Inhalt und Kontext extrahieren und deren semantische Bedeutung visualisieren (Wise et al. 1995). Neben gewöhnlichen digitalen Textgrundlagen reichen die Möglichkeiten von Klimateigenschaften bestimmter Regionen (Skupin & Esperbé 2008) über user-generierte Tags (Biberstine et al. 2010; Biberstine et al. 2012) bis zu online verfügbaren Sammlungen von Wikipedia-Artikeln (Bruggmann 2012; Salvini 2012). Ein Problem bei user-generierten Tags sind die Mehrdeutigkeiten einzelner Begriffe, die das Ergebnis verzerren können (Biberstine et al. 2012). Aus diesem Grund und anlässlich des diesjährigen *International Map Year* verwendet diese Arbeit einen Textkorpus bestehend aus Forschungsartikel der biennial

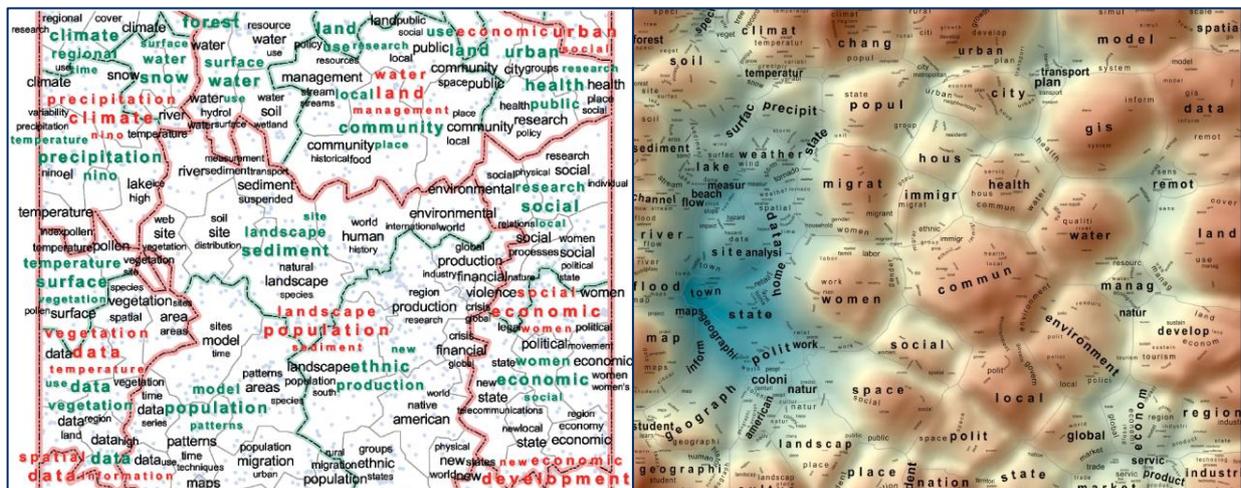


Abbildung 2-1: Geographische Informationslandschaften basierend auf Textdaten

stattfindenden *ICA*-Konferenzen (*International Cartographic Association*). Diese wurden entweder direkt über die offizielle Homepage der *ICA*<sup>3</sup> oder indirekt über die Verbindung zu den einzelnen Konferenz-Webseiten bezogen. Der zeitliche Rahmen reicht also von der Konferenz in Köln (1993) bis zur Konferenz in Dresden (2013). Eine Informationsvisualisierung, basierend auf Abstracts, wurde schon von Skupin (2002a) gestaltet (Abbildung 2-1 links). Die Grundlage für diese Darstellung bilden 2'220 geographische Forschungsartikel von der Konferenz der *AAG* (*Association of American Geographers*) in Hawaii im Jahre 1999. Diese Forschungsartikel gehören auch zum Fundament einer weiteren generierten Landschaft (Abbildung 2-1 rechts) von Skupin (2010), in welcher 22'000 Artikel der alljährlichen *AAG*-Konferenzen zwischen 1993 und 2002 bearbeitet wurden. Während die *AAG* das ganze Spektrum der Geographie beinhaltet fokussiert sich die *ICA* auf den kartographischen Forschungszweig der Geographie. In den Visualisierungen von Fabrikant & Salvini (2011), welche 2'693 *ICA*-Konferenzartikel aus den Jahren 1999 bis 2009 für eine Verräumlichung verwendet haben, sieht man bedeutende Schwerpunkte der kartographischen Wissenschaft. Zusätzlich zeigen die verschiedenen Kommissionen der *ICA* die Vielfalt der kartographischen Forschung. Während bei der Arbeit von Fabrikant & Salvini (2011) die Bildung der Regionen im Fokus stehen beschäftigten sich Skupin & de Jongh (2005) mit der Kontinuität der kartographischen Landschaft. Neben den bereits beschriebenen Arbeiten gibt es noch weitere generierte Informationslandschaften, die auf Fliesstextdaten wie Titel oder Abstracts basieren (Skupin & Buttenfield 1996; Skupin & Buttenfield 1997; Skupin et al. 2013; Skupin 2004).

Dieser Vielfalt an Datengrundlagen und Informationslandschaften liegen ebenso vielseitige Methoden zugrunde. Doch nicht-strukturierte Textdaten müssen unabhängig von der Methodik zunächst geordnet und die bedeutenden Elemente aus den Texten sowie Metadaten für weitere Analysen extrahiert werden, damit man Abstandsmessungen und andere Berechnungen an ihnen vornehmen kann. Ein Schritt, um Textdaten wie Konferenzartikel zu strukturieren, ist es, jeden einzelnen in getrennte Attribute wie «Jahr», «Autor(en)», «Titel», «Abstract» und «Ort der Konferenz» zu unterteilen (Skupin & Fabrikant 2003). Auf diese Weise resultieren semi-strukturierte Daten einerseits mit Attributen, die für die direkte weitere Analyse verwendet werden, andererseits mit weiteren Merkmalen, die als Metadaten zur Nachvollziehung der einzelnen Bearbeitungsschritte genutzt werden. Die notwendigen analytischen Ansätze hierfür reichen von klassischen Techniken des *Information Retrieval* (Jessup et al. 1999) bis zur automatischen Klassierung von extrahierten Texten (Chen et al. 1994).

Das Attribut «Jahr» in den beschriebenen Metadaten dieser semi-strukturierten Daten weist auf die vorhandene zeitliche Dimension hin. Es gibt nun verschiedene Möglichkeiten, diese darzustellen. In dieser Arbeit werden neben den Informationslandschaften auch Möglichkeiten der temporalen Darstellung solcher nicht-räumlicher Daten aufgezeigt. Ein Orientierungspunkt dabei sind Arbeiten von Skupin & Hagelman (2005) oder Skupin (2009). Eine Möglichkeit der Transformation von Textinhalt in eine räumliche Abbildung ohne Verlust der Dateneigenschaften ist im nächsten Unterkapitel mit der entsprechenden Technik beschrieben.

---

<sup>3</sup> <http://icaci.org/publications/> (Letzter Zugriff: 25.08.2015)

### 2.1.2 DIE VERRÄUMLICHUNG

Um die nicht-räumliche Datengrundlage interpretierbar zu visualisieren benötigt man ein Konzept, welches die Informationen strukturiert und den Eigenheiten der Information entsprechend darstellt. Das Wort «Information» bedeutet den Gehalt aus einer Ansammlung von Fakten (Skupin & Buttenfield 1997). Unabhängig von der Fähigkeit und dem Willen der betrachtenden Person, einen Sinn in diesem Set zu erkennen, ist die Information im Gegensatz zu Daten weder in einem Informationssystemen enthalten, noch wird sie von so einem übermittelt. Leute extrahieren die Information aus vielfältigen grafischen, textuellen und numerischen Ergebnissen solcher Informationseinrichtungen. Diese bilden ein Zeichensystem, welches gezielt für die Bedeutung entziffert werden muss (Couclelis 1998). Ein Informationssystem enthält und liefert keine Informationen. Es erlaubt dem Nutzer oder der Nutzerin, die Information kognitiv aus der Beziehung zwischen den Outputs des Systems und den eigenen aktuellen Interessen und früheren Erfahrungen zu konstruieren. Die Absicht von Informationsvisualisierungen ist die Vereinfachung der kognitiven Aufgabe, mit dem Nutzen des Verstands einem visuellen Muster eine Bedeutung zuzuweisen. Es handelt sich bei solchen Visualisierungen oft um niedrig-dimensionale Projektionen von hoch-dimensionalen Informationsräumen, welche aus der Gesamtheit von zweckgebundenen Elementen abstrahiert und konstruiert werden (Couclelis 1998). Ein Informationsraum ist eine graphische Darstellung, die basierend auf einer räumlichen Metapher und mit angewandten, alltäglichen Konzepten wie Distanz und Anordnung arbeitet (Fabrikant 2000). Diese Informationsräume reorganisieren so den Datenbankinhalt und stellen diesen basierend auf einer räumlichen oder einer geographischen Metapher dar (Fabrikant & Buttenfield 2001).

Während die Visualisierung der Welt mit georeferenzierten Daten seit Jahrhunderten ein kartografisches Anliegen ist (MacEachren & Kraak 2001), können die in der Erforschung räumlicher Informationen angewandten Metaphern und kognitiven Fähigkeiten in der nicht-räumlichen Informationsvisualisierung ebenso wertvoll sein (Skupin 2000). Denn die menschliche Vorstellungskraft in Kombination mit fachlichem Wissen sind leistungsfähige Werkzeuge, die zusammen mit Berechnungstools grosse heterogene Datenmengen in Informationen und anschliessend in Wissen transformieren können. Neben georeferenzierten Daten kann man auch in der Informationsvisualisierung eine kartographische Lösung suchen. Obwohl die räumliche Referenz durch ein Attribut eine Verknüpfung zu irgendeinem Ort aufweist in der Visualisierung nicht-räumlicher Informationen fehlt, lassen sich diese Daten in einer Verräumlichung in einem zwei- oder dreidimensionalen Raum, dem Informationsraum, und analog zu räumlichen Informationen anhand den semantischen Beziehungen in den Daten darstellen. Das Ziel dabei ist die effektive Darstellung nicht-räumlicher Daten in einer kognitiv verträglichen Art (Moore & Bricker 2015). Wie in Tabelle 2-1 zu erkennen ist, verleiht man der nicht-räumlichen Information eine räumliche Komponente als Fundament der Verräumlichung (de Jongh 2003). Beispiele solcher Darstellungen sind in Abbildung 2-1 zu erkennen.

Tabelle 2-1: Grobes Prinzip der Verräumlichung

Information	Prozess	Visualisierung
Keine räumliche Information		Graph, Diagramm, Abbildung
Räumliche Information		Verräumlichung Karte, Graph

Quelle: de Jongh (2003)

Dabei sollen Karten dieser Informationen visuell das Denken über die Beziehungen und räumlichen Muster ankurbeln (Kraak 2003). Doch wie bei gewöhnlichen Karten vereinfacht eine erfolgreiche Zuordnung von geographischen Konzepten zu spezifischen räumlichen Konfigurationen in der Darstellung die Interpretation solcher Beziehungen und räumlichen Muster. Aber auch die Vertrautheit der betrachtenden Person mit solchen Karten oder angedeuteten Landschaften beeinflusst die Wirkung der Verräumlichung (Moore & Bricker 2015). Vor allem die fundamentalsten geographischen Konzepte müssen direkt und intuitiv von der Repräsentation ablesbar sein. Zu diesen Konzepten gehören zum Beispiel der Ort oder Punkt, der Weg oder die Linie und die Region oder Fläche (National Research Council 1997). Der Ort ist die absolute und relative Positionierung von geometrischen Objekten im Informationsraum und wird empirisch als Punkt dargestellt. Bei einer solchen Verräumlichung ist die relative Position eines Datenpunkts gleichzeitig eine Repräsentation der Beziehungen innerhalb des Datensatzes (Chalmers 1993). Diese Positionierung erfolgt analog zum ersten Gesetz der Geographie nach Tobler (1970), nach welchem Leute annehmen, dass nähere Punkte auch einander ähnlicher sind. Diesen Grundsatz nennt man Distanz-Ähnlichkeits-Metapher (Montello et al. 2003). Der Weg oder die Grenze repräsentieren Linien, welche Orte im Informationsraum entweder verbinden oder trennen. Die Grenze teilt die Daten in Regionen ein, die anschaulich als Fläche dargestellt werden und so die Orte im Informationsraum ordnen und einteilen. Die Grenzen zwischen den Regionen sind in der Informationsvisualisierung nicht unbedingt geschlossen, permanent, homogen oder kontinuierlich (Couclelis 1998). Durch das Trennen und Sammeln von Punkten als hierarchisches Schema wird die Bedeutungszuweisung einzelner Punkte vereinfacht und Gemeinsamkeiten oder Differenzen verdeutlicht (Friedman & Brown 2000). Die Datenpunkte sind eine geeignete Grundlage in der Landschaftsvisualisierung. So besteht die Möglichkeit ein geeignetes Attribut diesen Punkten zuzuweisen und dann daraus eine kontinuierliche Informationsoberfläche zu interpolieren (Wise et al. 1995).

Der Informationsraum bietet mithilfe dieser geographischen Konzepte eine klar abgegrenzte Strategie zum Organisieren von Informationen. So kann er mit logischen, gedächtnishelfenden, metrischen oder nicht-metrischen Koordinaten definiert werden. Damit kann man die Navigation, Übersicht und auch die Datenabfragen in diesem generischem Informationsraum vereinfachen. Es handelt sich um ein aus den Daten generisch erstelltes Gerüst basierend auf geographischen Metaphern, um den Zugriff auf den Inhalt der Informationen zu unterstützen (Skupin & Buttenfield 1997). Dabei spielen die genannten geographischen Metaphern das Grundgerüst dieser Struktur. Durch das Generieren einer sinnvollen synthetischen räumlichen Struktur (wie es zum Beispiel die Distanz zwischen einzelnen Objekten ist) werden also enorme Mengen an abstrakten Daten, welche gewöhnlich aus einem Textkorpus stammen, in eine verständlichere und visuell kompakte Form gerendert und mittels Gestaltungskonzepten der topographischen und der thematischen Kartographie dargestellt. (Dodge & Kitchin 2001)

Gemäss Skupin & Fabrikant (2003) gibt es innerhalb der Verräumlichung verschiedene Methoden, die Daten in den künstlich erstellten Informationsraum zu projizieren und visualisieren. So beschreibt Skupin

(2002b) die Möglichkeiten analog zur geometrischen Repräsentation der Information. Diskret kann man die Information als Punkt (*MDS & PCA*), als Linie (Kurvendiagramme & die H-Baum-Anordnung) oder als Fläche (Baumkarte) darstellen. Beispiele dazu sind in der Arbeit von Skupin & Fabrikant (2003) zu finden. Basierend auf einem Feld-Konzept, welches für die kontinuierliche, aus den Punktdaten interpolierte, Informations-oberfläche benötigt wird, wird die Technik der *SOM* angewandt (Kohonen 2001). Auch de Jongh (2003) schlägt in seiner Arbeit neben der *MDS* die *SOM*-Technik als Verräumlichungsprozess vor. Diesen kann man in drei Schritte unterteilen. In Abbildung 2-2 wird der erste Prozess (1) für das Sammeln, Analysieren und Prozessieren der Daten veranschaulicht. Danach (2) folgt die erste Verräumlichung der Daten in einem konzeptuellen Informationsraum. Die Anwendung von kartographischen Methoden und Techniken in der

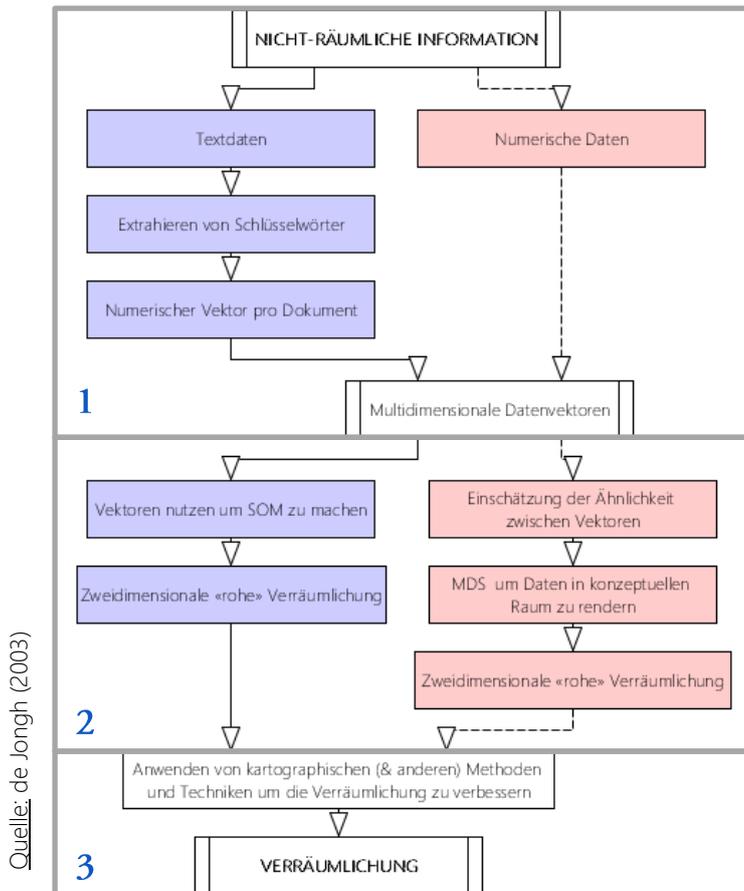


Abbildung 2-2: Verfeinertes Prinzip der Verräumlichung

Verräumlichung (3) lässt die ursprünglich nicht-räumlichen Informationen wie räumlich referenzierte Daten erscheinen. Die durchgehenden Pfeile symbolisieren den in dieser Arbeit angewandten Verräumlichungsprozess. Die gestrichelten Pfeile zeigen einen alternativen Prozess. (de Jongh 2003)

Mit dem Feld-Konzept, das einer *SOM* zugrunde liegt, lässt sich eine Informationslandschaft aus den Eingangsdaten erodieren. Diese zählt als eine der ersten Methoden, welche in der Informationsvisualisierung verwendet wurde, um Datenarchive zu reorganisieren und darzustellen (Fabrikant et al. 2010). Solche Landschaften kann man analog zu gewöhnlichen Karten wie der topographischen oder thematischen Karte gestalten. Auch das hierarchische Schema aus den Karten kann man übernehmen, bei welchen die Städte in Ländern liegen und sich die Länder wiederum auf verschiedenen Kontinenten ausbreiten (Friedman & Brown 2000). Neben den beschriebenen Konzepten steht in solchen virtuellen Informationslandschaften die räumliche Abstandsmetapher im Mittelpunkt. Mit der relativen Positionierung der Grundlagedaten im Informationsraum repräsentiert man die semantische Ähnlichkeit zwischen Datenelementen in sehr grossen Datenbanken (Card et al. 1999). Diese Konzeptualisierung und Formalisierung der räumlichen Nähe ist noch kritischer, wenn nicht-räumliche Daten in ein logisch, aber beliebig definiertes räumliches Bezugssystem, basierend auf der semantischen Ähnlichkeit, projiziert werden (Fabrikant et al. 2002). Die Darstellungen dieser Arbeit orientieren sich an der Übertragung der Darstellungsprinzipien einer politisch gegliederten und einer Relief-Karte auf die nicht-geographische Informationslandschaft.

### 2.1.3 SELF-ORGANIZING-MAPS (SOM)

Eine Visualisierungstechnik, welche nicht explizit räumliche Informationssammlungen räumlich anordnet und darstellt, ist als Verräumlichung benannt (Fabrikant et al. 2010). Die grösste Herausforderung in dieser Technik ist das Reduzieren der Datendimensionen. Die Schwierigkeit des Ableitens einer geeigneten, niedrig-dimensionalen geometrischen Konfiguration aus einem hoch-dimensionalen Datensatz erhöht sich bei einer Zunahme der beteiligten Dimensionen im Transformationsprozess (Skupin & Fabrikant 2003).

Wie schon in Abbildung 2-2 zu sehen ist werden in dieser Arbeit die Methode der *SOM* verwendet. Diese reduziert die Datendimensionen und stellt diese in einem zweidimensionalen Informationsraum visuell ansprechend dar (Börner et al. 2003). Eine andere Methode ist die *MDS* (Multidimensionale Skalierung). Im Gegensatz zur *MDS* fokussieren sich *SOM* mehr auf den Erhalt der topologischen Beziehungen als auf die anfänglichen Ähnlichkeitseigenschaften der entfernteren Beziehungen (Skupin & Fabrikant 2003). Bei dieser künstlichen, neuronalen Netzwerkmethod werden multidimensionale Eingangsdaten in einen zweidimensionalen topologischen Informationsraum projiziert. Dieser besteht bei *SOM* aus einer rechteckigen oder hexagonalen Parzellierung an Neuronen (Kohonen et al. 2000). Es ist ausserdem auch die häufigste Informationsvisualisierungsmethode, welche auf einem Feld-Konzept basiert (Kohonen 2001). Dieses Feld-Konzept ist vor allem für die Gestaltung einer Informationslandschaft mit einer kontinuierlichen Informationsoberfläche ähnlich einer Relief-Karte bedeutsam.

Neben der Dimensionsreduktion, die bei der Projektion von n-dimensionalen Daten in einen zweidimensionalen Informationsraum stattfindet, können *SOM* auch eine Clustertechnik einschließen. Dabei dienen die Neuronen an sich als ein Cluster, welche verschieden viele Grundlagedaten beinhalten (Boyack et al. 2011). Und je nach Wahl der Anzahl Neuronen müssen die Inputdaten auf weniger oder mehr Neuronen als Vektoren verteilt werden, was zu einer automatisierten Clusterung des Datensatzes führen kann (Skupin & Agarwal 2008). Die Anzahl Neuronen kann man mit der Auflösung eines Satellitenbildes vergleichen (Fabrikant & Skupin 2005). Neben den Neuronen, welche das Fundament der Umgebung des Informationsraums bilden, nutzen *SOM* eine weitere Komponente, die *best-matching unit (BMU)*. Die *BMU* ist die Repräsentation eines Datenelements als Punkt in dieser Neuronenumgebung. Dabei werden diese *BMU* in das Neuron projiziert, welches die optimalsten Eigenschaften für den jeweiligen Eingangsdatenpunkt aufweist. So erhält jedes einzelne Datenelement des Datensatzes ein Neuron zugewiesen, wobei ein Neuron auch mehrere *BMU* enthalten kann. Die Neuronen hingegen nutzen die sogenannte U-Matrix, welche die n-dimensionalen Differenzen benachbarter Neuronen beziffert. Die Neuronen übernehmen zudem die gleichen Attribute wie die Inputdaten (Skupin & Agrawal 2008). Dies wiederum erlaubt eine erste Prüfung des Resultats, indem die räumliche Verteilung der Neuronen mit ähnlichen Attributen analysiert wird.

Trotz der punktuellen Projektion von diskreten Objekten im Kontinuum der Informationslandschaft bieten *SOM* ganz verschiedene Visualisierungsmöglichkeiten (Skupin & Esperbé 2008). Wie in der Einleitung beschrieben ist das Ziel der Arbeit, thematische und topographische «Karten» solcher nicht-räumlicher Informationen zu gestalten.

#### 2.1.4 DISKRET VS. KONTINUIERLICH

Das Feld-Konzept, auf welchem die *SOM*-Methode gestützt ist (Kohonen 2001), ermöglicht es, die Visualisierung ebenso kontinuierlich oder diskretisiert zu gestalten, wie beim geographischen Raum, welcher diskrete Objekte und kontinuierliche Felder beinhaltet, die kombiniert oder getrennt dargestellt werden können. So handelt es sich bei den kontinuierlichen Rasterdaten in Informationsvisualisierungen oft auch um Bilddaten und Interpolationen räumlich-gestreuter Daten. Die diskreten sind entweder klassierte Daten oder exakte Positionen einzelner Komponenten (Tory & Moller 2004). Während die diskreten Objekte als geometrische Vektordaten wie Punkte, Linien oder Polygone abgebildet werden, füllen die kontinuierlichen Rasterdaten den gesamten Raum mit Informationen. Diese Informationen sind keine eigenständigen Objekte, sondern voneinander abhängig. Dennoch lässt sich die Landschaft durch Täler und Hügel als Felder darstellen. Die Vektoren werden als Punkte, Linien oder Polygone dargestellt. Das beste Beispiel eines diskreten Elements sind Grenzen. Das Problem dabei ist, wo man kontinuierliche Phänomene wie zum Beispiel ein Tal abgrenzt. Es handelt sich dabei um eine empirische Fragestellung. Sowohl kontinuierliche Phänomene als auch scharf abgegrenzte Elemente in der Landschaft können entsprechend visualisiert werden. Die flüchtige Art der Felder steht im starken Kontrast mit der starken Individualität der lokalisierten und individuellen Objekte im vektor-basierten Konzept (Couclelis 1992). Ausserdem erklärt Rosch (1973), dass, wenn Menschen grosse Schwierigkeiten mit kontinuierlichen Phänomenen haben, die diskrete Repräsentation der kontinuierlichen Eigenschaften des Raumes kognitiv angebracht ist, da die Menschen auf diese Weise einen Sinn in der Umwelt erkennen.

Bei nicht-geographischen Informationsvisualisierungen spielt dieser Unterschied auch eine bestimmende Rolle. Entscheidet man sich für eine diskrete Konzeptualisierung der Informationen kann die Visualisierung als Netzwerkvisualisierung oder auch als *MDS* gestaltet werden. Bei einer kontinuierlichen Konzeptualisierung bietet die *SOM* eine Möglichkeit, die Informationen als Kontinuum zu analysieren. (Fabrikant & Skupin 2005)

Bei nicht-geographischen Informationen wie Konferenzbeiträgen stellt die Wahl der angemessenen Konzeptualisierung eine Herausforderung dar. Skupin (2009) beschreibt die damit verbundenen Problematiken in Zusammenhang mit einer Informationsvisualisierung der geographischen Wissenschaft. Dabei wird die zeitliche Entwicklung einzelner Autoren in der Informationslandschaft als diskrete Trajektorie und als eine kontinuierliche Änderungsfläche visualisiert. Bei der letzten Darstellung zeigen verringerte oder erhöhte Übereinstimmungen analog einer Art *Heatmap* die zeitliche Entwicklung eines Autors. Die diskrete Lösung eines Pfades entlang der Zeitlinie wird auch in dieser Arbeit als Nebenprodukt erstellt.

Bei Informationsvisualisierungen ist der Unterschied zwischen der diskreten und der kontinuierlichen Konzeptualisierung anhand geographischer Metaphern wie Distanz oder Region noch unerforscht. Der Visualisierungstyp K in dieser Arbeit stellt die Oberfläche der Informationslandschaft kontinuierlich dar. Ihr Erscheinen orientiert sich an einer Relief-Karte. Die Oberfläche der anderen Informationslandschaft ist als Gegensatz zu einer Relief-Karte klassiert und ist vergleichbar mit einer politisch-gegliederten Karte (Visualisierungstyp D) mit Höheninformationen.

## 2.2 VISUALISIERUNGSBEISPIELE UND VERÖFFENTLICHTE STUDIEN

Neben der methodischen Grundlage sind Visualisierungsbeispiele und bisherig veröffentlichte Studien bedeutend. Die Visualisierungsbeispiele dienen als Vorstellung, wie die zu vergleichenden Darstellungen aussehen könnten und als methodisches Vorbild, da einzelne Arbeitsschritte dadurch erklärt und begründet werden. Ausserdem geben die Beispiele auch einen Eindruck in die Vielfalt der Visualisierungen, die mit dieser Methodik gestaltet wurden. Die Resultate bereits veröffentlichter Studien sind wichtig, da sie schon einen Einblick in mögliche Ergebnisse der im Zusammenhang mit dieser Arbeit durchgeführten Studie geben können und so mögliche Hypothesen und Fragestellungen begründen. Dabei ist zu erwähnen, dass nur Visualisierungsbeispiele oder veröffentlichte Studien besprochen werden, welche für einen Vergleich zweier Informationslandschaften relevant sein können.

### 2.2.1 VISUALISIERUNGSBEISPIELE

Während eine der beiden Endvisualisierungen eher die kontinuierliche Informationsoberfläche wie in einer topographischen Karte betonen soll fokussiert sich die andere Darstellung auf die klassierte Anordnung der Informationen. Dazu sind schon verschiedene Informationslandschaften gestaltet worden, bei welchen man sich über eine der beiden Möglichkeiten orientieren kann. So haben Biberstine et al. (2010; 2012) eine kontinuierliche (Abbildung 2-3: links) und eine diskretisierte (Abbildung 2-3: rechts) Musiklandschaft visualisiert. Bei beiden sind die Inputdaten usergenerierte Tags des Online-Musikkatalogs Last.Fm<sup>4</sup>. Trotz der gleichen Datengrundlage bietet die Visualisierung links mit der kontinuierlich dargestellten Variablen der Verbindungsstärke des häufigsten Musiktags zum Neuron eine Mehrinformation.

Beide Musiklandschaften zusammen ergeben einen guten Eindruck der darstellerischen Möglichkeiten und zeigen auch *SOM*, die nur diskretisiert oder kombiniert dargestellt werden.

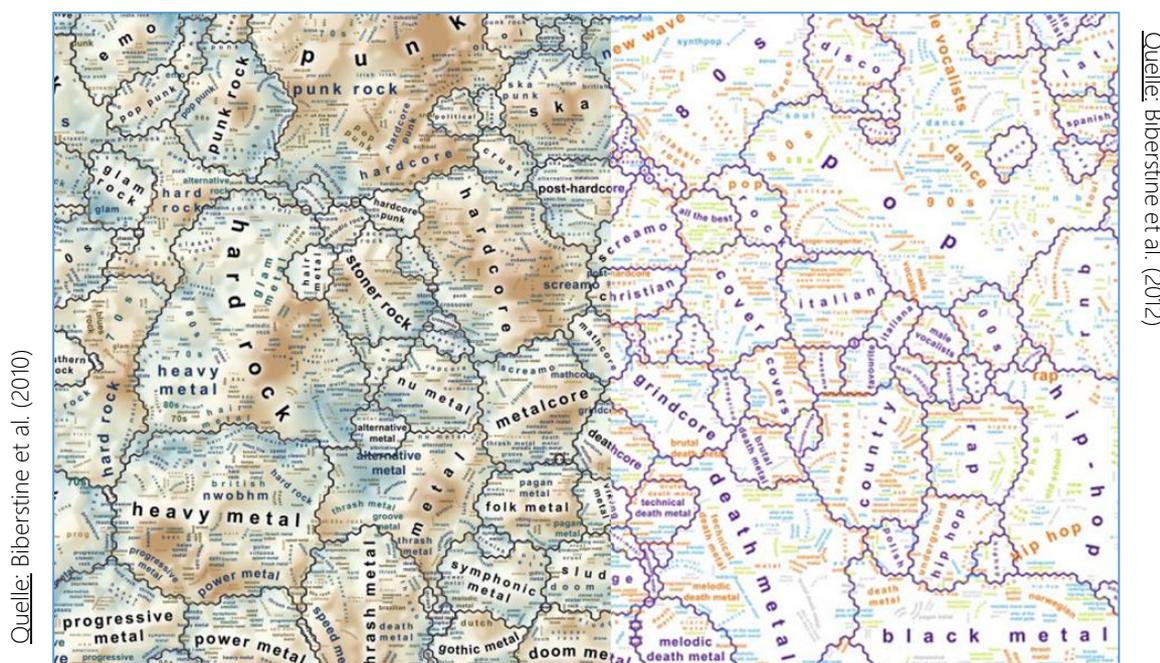


Abbildung 2-3: Verschiedene Darstellungen einer Musiklandschaft

<sup>4</sup> <http://www.last.fm/> (Letzter Zugriff: 12.09.2015)

Neben der Musiklandschaft hat sich Skupin (2002a, 2004) mit Visualisierungen der geographischen Wissenschaft beschäftigt, die er einmal wie im Beispiel kombiniert (Abbildung 2-4: links) und diskretisiert (Abbildung 2-4: rechts) darstellt. Bei beiden Darstellungen bilden Konferenzbeiträge der AAG die Grundlage für die Informationslandschaften. Als Variable für die kontinuierliche Oberfläche wird dabei die Begriffsdominanz, welche den Vorrang der drei häufigsten Begriffe pro Neuron ausdrückt, verwendet. Hingegen sind bei der diskretisierten Visualisierung neben den Grenzen der einzelnen Regionen auch die einzelnen Konferenzbeiträge als graue Punkte dargestellt.

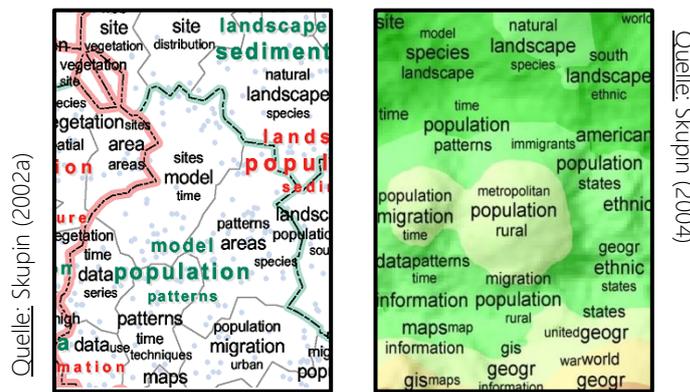


Abbildung 2-4: Geographische Landschaft basierend auf AAG-Konferenzbeiträge

Eine in den Visualisierungen erkennbare Problematik ist die Aggregation von Punkten zu Regionen. Man kann dabei, wie es in beiden Arbeiten auch gemacht wird, die Anzahl der Cluster variieren. Während die diskretisierte Variante mit einer unterschiedlichen Anzahl Regionen arbeitet, nutzt die kontinuierliche Darstellung verschiedene Beschriftungsebenen. Die geographische Wissenschaft wird auch in weiteren Arbeiten als Datengrundlage verwendet. Zum Beispiel basiert die Informationslandschaft von Fabrikant & Salvini (2011) auf der ICA. Dabei wurden die Konferenzbeiträge von 1999 (Ottawa) bis 2009 (Santiago) zuerst in einem Netzwerk und anschließend als «Karte» visualisiert, in welcher die diskrete Anordnung der einzelnen kartographischen Cluster im Vordergrund steht. Dieser Artikel dient bei einem Grossteil des Arbeitsprozesses einer solchen Visualisierung, inklusive Gruppierung der Konferenzbeiträge, als Vorlage. Ausserdem konnte für diesen Zeitrahmen die bereits aufbereitete semi-strukturierte Datengrundlage von Fabrikant & Salvini (2011) in dieser Arbeit verwendet werden.

Zusätzlich wurde von Skupin & de Jongh (2005) eine ICA-basierte, kontinuierliche Landschaft gestaltet. Als Grundlage der Abbildung 2-5 dienen die beiden ICA-Konferenzen in Peking (2001) und Durban (2003). Die als Topographie dargestellte Begriffsdominanz gibt zusammen mit den zu Wortstämmen gekürzten

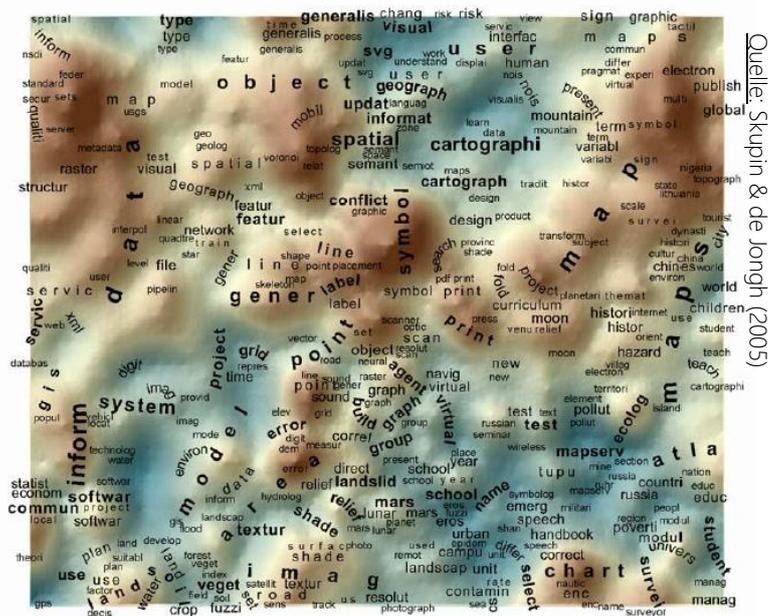


Abbildung 2-5: ICA-Landschaft der Konferenzen in Peking (2001) und Durban (2003)



### 2.2.2 VERÖFFENTLICHTE STUDIEN

Eine für diese Arbeit bedeutende Publikation ist von Fabrikant et al. (2006). In ihrer Arbeit wurde der Einfluss verschiedener geographischer Metaphern wie Distanz und Region auf die relative Ähnlichkeitswahrnehmung von Punkten in einem nicht-geographischen Informationsraum untersucht. Dabei stellte sich heraus, dass neben der Distanz und Region auch die Gruppierung und die Anordnung der Punkte einen Effekt auf die Wahrnehmung der Ähnlichkeit haben. Des Weiteren kann man mit der Farbgebung der Räume den Einfluss der geographischen Metapher «Region» verstärken oder mindern. So werden Räume, die mit einem ähnlichen Farbton gestaltet sind, auch semantisch als ähnlicher betrachtet. Zudem resultiert aus der Studie, dass die Region die Ähnlichkeitswahrnehmung mehr beeinflusst als die räumliche Distanz. Bei der bearbeiteten Darstellung in dieser Studie handelt es sich um eine für diesen Fragentyp entwickelte Anordnung von Punkten und Polygonen. Unter anderem mit diesem Aufgabentyp sollen in dieser Arbeit die beiden Informationslandschaften verglichen werden. Dabei sollen diese Informationslandschaften mit einer zusätzlichen Variablen die Morphologie der Informationen darstellen.

Dazu wurde schon untersucht, ob die Landschaftsmetapher einen Einfluss auf die Wahrnehmung der Morphologie der Informationslandschaft hat. Die Resultate dieser veröffentlichten empirischen Studie haben ergeben, dass die Verständlichkeit der Morphologie nicht so intuitiv ist wie erwartet und dass ungeübte Anwender und Anwenderinnen Schwierigkeiten haben, diese räumliche Metapher vollständig zu erfassen und zu interpretieren. Begründet wird es mit einem Mangel an Vertrautheit einer solchen Darstellung sowie mit einem allgemein laienhaften Verständnis der Geomorphologie (Fabrikant et al. 2010). Neben der Vertrautheit spielen auch andere Fähigkeiten wie die Raumvorstellung eine Rolle im Verständnis von gewöhnlichen Karten (Ishikawa & Kastens 2005). Aus diesem Grund möchte diese Studie einen Unterschied zwischen den Geographie-Studentinnen und -Studenten und Personen mit einer anderen Ausbildung untersuchen.



### 3 FORSCHUNGSFRAGEN UND METHODIK

#### 3.1 ARBEITSGIELE

Am Anfang dieses Kapitels sind die wichtigsten Ziele erklärt, die mit dieser Arbeit erreicht werden sollen. Die gesamte Arbeit teilt sich auf in das Arbeitsziel I, welches das Erstellen gewünschter Visualisierungen beinhaltet, und das Arbeitsziel II, das diese Visualisierungen in einer Studie auf deren Verständlichkeit untersucht. Dieses aufeinander aufbauende Prinzip ist in Abbildung 3-1 zu sehen.

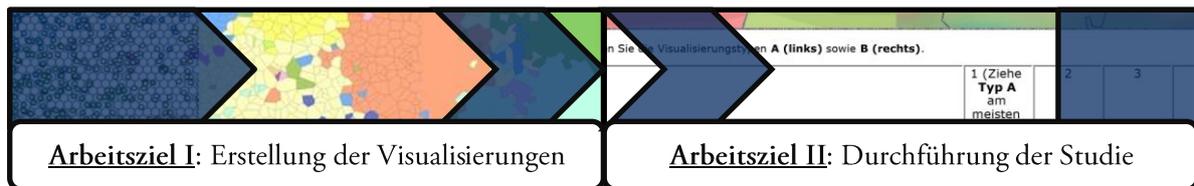


Abbildung 3-1: Arbeitsziele

Zur Bearbeitung des ersten Arbeitsziels werden keine neuen Methoden entwickelt, sondern auf die im vorherigen Kapitel vorgestellten und bereits etablierten Methoden zurückgegriffen. Die Kombination dieser Methoden und der genaue Visualisierungsprozess mit allen einzelnen Arbeitsschritten ist in Kapitel 3.4 beschrieben und begründet. Wie schon in der Einleitung beschrieben soll das Design der Informationslandschaften den politisch-gegliederten und den Relief-Karten gleichen. Das zweite Arbeitsziel setzt sich aus der Vorbereitung, Durchführung und Auswertung einer gesamten Studie zusammen. Diese Studie basiert auf den im ersten Arbeitsziel erstellten Visualisierungen und soll die im Unterkapitel 3.2 formulierten Hypothesen und Fragestellungen untersuchen.

## 3.2 FRAGESTELLUNGEN & HYPOTHESEN

Aus den dargelegten Arbeitszielen lässt sich eine Vielzahl an möglichen Fragestellungen ableiten. Diese Arbeit legt den Schwerpunkt auf den Ausschnitt, welcher sich mit dem Verständnis von den im ersten Arbeitsziel erstellten Informationslandschaften beschäftigt. Der Arbeitsprozess zu solchen Informationslandschaften ist bekannt und erfordert keine separaten Forschungsfragen. Die detaillierten Fragestellungen sowie die ausformulierten Hypothesen zu den jeweiligen Fragestellungen sind in diesem Unterkapitel erläutert.

### 3.2.1 FRAGESTELLUNGEN

Wie man aus den vorgestellten Visualisierungsbeispielen erkennen kann, gibt es viele verschiedene Möglichkeiten, eine solche Informationslandschaft darzustellen. Diese Arbeit vergleicht zwei Beispiele, angelehnt an einer politischen und einer Relief-Karte, die speziell für diesen Vergleich angefertigt werden. Die Leitfrage ist, ob das Übertragen der geographischen Metaphern wie Region oder Distanz bei beiden Visualisierungstypen gleich funktioniert. Dabei kann sie in mehrere Forschungsfragen unterteilt werden, welche das Studiendesign prägen (Tabelle 3-1).

Tabelle 3-1: Forschungsfragen

Leistung	<i>Performance</i>	«Beeinflusst der Visualisierungstyp die Leistung beim Beantworten der gestellten Aufgaben?»
	<i>Sicherheit</i>	«Beeinflusst der Visualisierungstyp das Sicherheitsgefühl beim Beantworten der gestellten Aufgaben?»
Präferenz		«Präferieren die Testpersonen eine der beiden Informationslandschaften abhängig von der gestellten Aufgaben?»
Hintergrundwissen		«Beeinflusst der Grad der geographischen Ausbildung die Leistung beim Beantworten der gestellten Aufgaben?»

### 3.2.2 HYPOTHESEN

Abgeleitet aus den Forschungsfragen finden sich in diesem Kapitel die dazu formulierten Hypothesen (Tabelle 3-2), welche in dieser Arbeit untersucht werden. In Kapitel 5 werden die Hypothesen nochmals aufgegriffen und abhängig von den Resultaten dieser Arbeit gezeigt, ob diese abgelehnt oder angenommen werden.

Tabelle 3-2: Hypothesen

H <sub>(1.1)</sub> :	Der Visualisierungstyp beeinflusst die Leistung beim Beantworten der gestellten Aufgaben.
H <sub>(1.2)</sub> :	Der Visualisierungstyp beeinflusst das Sicherheitsgefühl beim Beantworten der gestellten Aufgaben.
H <sub>(2)</sub> :	Die Testpersonen präferieren einen Visualisierungstyp abhängig von der gestellten Aufgabe.
H <sub>(3)</sub> :	Eine geographische Ausbildung beeinflusst die Leistung beim Beantworten der gestellten Aufgaben.

### 3.3 DATEN: GRUNDLAGE & AUFBEREITUNG

Wie bei Fabrikant & Salvini (2011) bilden in dieser Arbeit die Titel und Abstracts von Forschungsartikeln der biennial stattfindenden *ICA*-Konferenzen aus den Jahren 1993 bis 2013 die Datengrundlage. Verwendet wurden alle Artikel, welche auf der Homepage der *ICA*<sup>5</sup> zu finden sind und einen englischen Titel und/oder Abstract aufweisen. Dabei wurden neben den für die Visualisierung benötigten Text-Inputs wie Titel und Abstracts, die Autoren sowie das Jahr und der Austragungsort der Konferenz extrahiert. 2'693 bereits aufbereitete Konferenzbeiträge der Jahren 1999 bis 2009 konnten direkt von Fabrikant & Salvini (2011) übernommen werden, während die fehlenden selbst strukturiert und aufgearbeitet wurden. Insgesamt werden dabei 5'105 Artikel bearbeitet, von welchen aufgrund der Sprache nur 4'922 berücksichtigt werden können, da 147 Artikel in Französisch, 25 in Spanisch und 11 in Deutsch verfasst worden sind. Von diesen 4'922 weisen 4'502 einen Abstract auf, während die Übrigen nur mit einem englischen Titel beschrieben sind. Die Begründung dieser Datengrundlage ist im Kapitel 2.1.1 beschrieben. Um die Daten zu analysieren und zu visualisieren, müssen sie in eine passende Struktur gebracht werden. Als Ausgangsdatei gilt eine Vorlage, welche die nicht-strukturierten Konferenzbeiträge als eine semi-strukturierte Datensammlung anordnet (Skupin & Fabrikant 2003). Dies geschieht tabellarisch mit der Gliederung jedes einzelnen Konferenzbeitrags in Jahr, Autor, Titel, Abstract und jeweiligem Konferenzort. Um später eine Identifizierung der einzelnen Beiträge zu ermöglichen, wird jedem Beitrag auch eine eindeutige Identifikationsnummer (ID) zugewiesen. Als Beispiel der aufbereiteten Daten dient die Tabelle 3-3.

Tabelle 3-3: Semi-strukturierte Grundlagedaten

ID:	DATE:	AUTHORS:	TITLE:	ABSTRACT:	LOCATION:
1	1993	D.Rhind	Mapping for the new millenium	-	Cologne, Germany
...	...	...	...	...	...
4'922	2013	N. Jiang, Q. Sun, Y. Cao, H. Zhang, Y. Gu	Research and Application of Two-peak Changing Law of Electronic Map Load	Electronic map load is one of the most important factors that influence the clarity, smooth and gradation of the multi-scale display; and is also an important indicator in multi-scale map representation modeling. All of these make the research of electronic map load more significance. This article analyzed differences of the map load variation between paper map and electronic map, using RGB values-based approach to obtain a large number data of electronic maps, explored two-peak changing law of electronic map load by making further research; discussed the curve shape and influencing factors of two-peak changing law of electronic map load. In the end, established a more scientific and practical electronic map multiscale model, and has achieved a preferable experiment results.	Dresden, Germany

<sup>5</sup> <http://icaci.org/publications/> (Letzter Zugriff: 25.08.2015)

### 3.3.1 VERARBEITUNG DER DATEN (TOPIC MODELING)

Um die Artikel untereinander semantisch vergleichen zu können, gibt es verschiedene Möglichkeiten. In dieser Arbeit wird die Methode des *Probabilistic Topic Modeling* nach Steyvers & Griffiths (2007) angewandt. Diese Methode basiert auf der Idee, dass die zu bearbeitenden Textdokumente eine Mischung verschiedener Themen sind und gehört zur Klasse der latenten semantischen Analyseverfahren (Landauer et al. 1998). Ein Thema wiederum ist eine Wahrscheinlichkeitsverteilung einzelner Wörter. Ein thematisches Modell ist ein generatives Modell für Dokumente. Jedes in diesen Dokumenten vorkommende Wort exklusive den Stoppwörtern, die in einem späteren Abschnitt erklärt sind, wird in ein Thema oder mehrere Themen eingeteilt. Statistische Standardtechniken kehren diesen Prozess um und extrahieren Begriffe aus der Datengrundlage, welche für die Repräsentation einzelner Themen der Dokumentkollektion verantwortlich sind (Steyvers & Griffiths 2007).

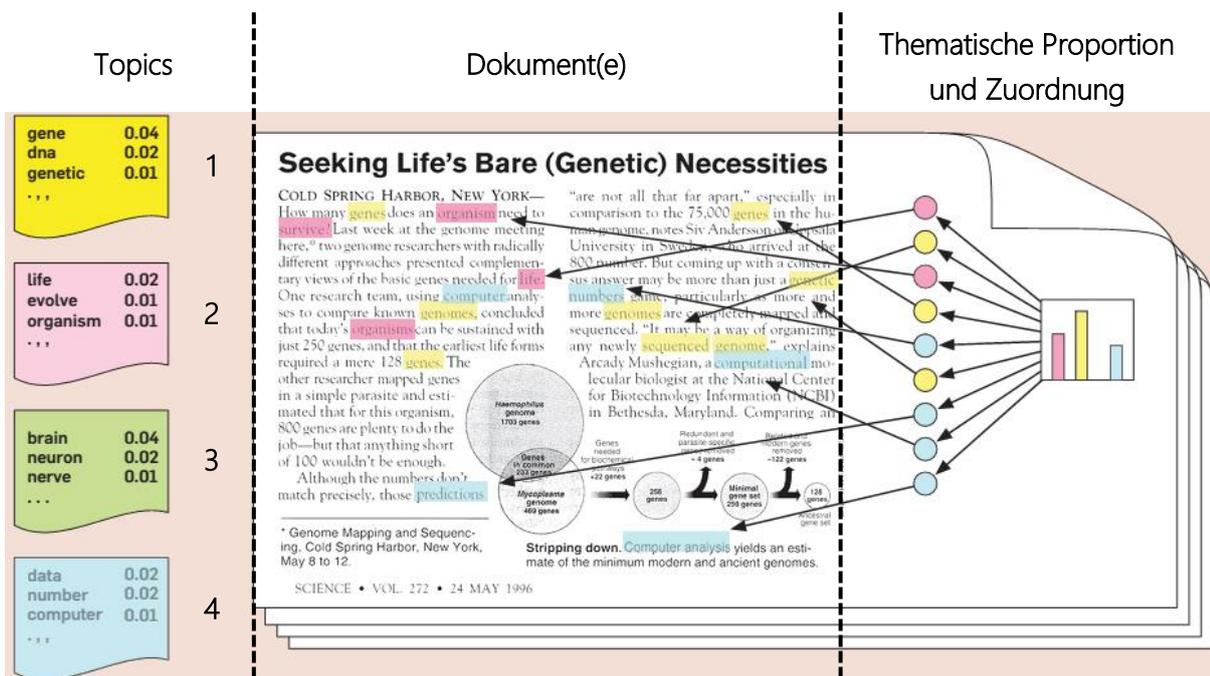


Abbildung 3-2: Schema des Topic Modeling

Das Vorgehen ist in Abbildung 3-2 schematisch mit vier Themen dargestellt. Die Themen ergeben sich aus der Wahrscheinlichkeit des gemeinsamen Vorkommens von Begriffen in einem Dokument der gesamten Datengrundlage. Das einzelne Dokument (hier: *Seeking Life's Bare (Genetic) Necessities*) wiederum wird anhand der Häufigkeiten dieser Begriffe der verschiedenen Themen gewichtet. In der Abbildung 3-2 erkennt man exemplarisch vier Themen inklusive der wahrscheinlichsten Schlüsselwörter und das Dokument, welches anhand dieser Schlüsselwörter in die generierten Topics eingeteilt ist. So sieht man die thematische Zuordnung des Dokuments, welches zum Beispiel die Identifikationsnummer 122 zugewiesen erhält (Tabelle 3-4). Aus der Zuordnung jedes einzelnen Objekts zu denselben in der Anzahl festgelegten, extrahierten Themen resultiert die Artikel-Themen-Matrix, welche die Grundlage für die Weiterverarbeitung der Daten bildet.

Tabelle 3-4: Vereinfachte exemplarische Artikel-Themen-Matrix I

ID	Topic 1	Topic 2	Topic 3	Topic 4
121	...	...	...	...
122	0.47	0.24	0.00	0.29
123	...	...	...	...

Überträgt man nun das Vorgehen für das oben dargelegte Beispiel auf die Datengrundlage dieser Arbeit, werden für das Topic Modeling die Kombination der Spalten «Titel» und «Abstract» der semi-strukturierten Konferenzbeiträge verwendet. Zusätzlich erlaubt eine durchgehende Nummerierung der Beiträge später eine eindeutige Identifizierung. Ausserdem sollen nicht vier wie im Beispiel sondern 20 Themen aus der Dokumentensammlung bestimmt werden. Diese scheinbar beliebige Anzahl lässt sich einerseits mit der Anzahl ICA-Kommissionen begründen, welche sich im Verlauf der Zeit zwischen 12 und 28 bewegt hat, andererseits stützt sich die Wahl in der Arbeit auf bisherige Publikationen (Fabrikant & Salvini 2011; Bruggmann et al. 2013a). Die vereinfachten resultierenden Themen dieser Arbeit und die dazugehörigen Schlüsselwörter sind in der Tabelle 3-5 ersichtlich. Die komplette Themenübersicht ist im (Anhang B) zu finden. Damit keine Themen bestehend aus Füllwörter wie «and» oder «how» zu entstehen, verwendet man eine englische Stopppwortliste. Diese Auflistung an Wörter wird in der Themengenerierung nicht berücksichtigt und schränkt das Risiko von unübersichtlichen Themen erheblich ein. Als Vorlage dient in dieser Arbeit die online verfügbare englische *NLTK*-Stopppwortliste<sup>6</sup> (Anhang A).

Tabelle 3-5: Vereinfachte Themenübersicht aus Topic Modeling

Topic	Schlüsselwörter	Topic	Schlüsselwörter
1	cartography, research, use, students	11	design, visualization, research, dynamic
2	process, quality, cartographic, means	12	information, use, context, people
3	geographical names, history, works, names	13	mapping, use, needs, techniques
4	image(s), topographic maps, accuracy, satellite images	14	environment, region, development, complex
5	model, method, based, generalization	15	different, approach, used, models
6	data, database, spatial data, spatial information	16	atlas, thematic maps, form, maps
7	map(s), map design, symbols, map use	17	project, development, production, standards
8	analysis, area(s), using, results	18	described, points, projection, fact
9	time, analysis, space, change	19	land use, classification, 3 d, use
10	users, internet, application, services	20	system, gis, information, technology

Für die automatisierte Berechnung einer Artikel-Themen-Matrix wie sie als Beispiel in Tabelle 3-4 zu sehen ist sowie für die Extraktion der Themen wird die Text-Visualisierungs-Toolbox (TVT) in MATLAB (Mathworks 2010) genutzt (Hespanha & Hespanha 2011). Diese Toolbox beinhaltet die MATLAB Topic Modeling Toolbox (1.3.2) mit der Implementierung der Latent Dirichlet Allocation, die von Steyvers & Griffiths (2007) entwickelt wurde.

Vergleicht man nun die in diese Toolbox eingehende mit der ausgehenden Datenmenge fällt auf, dass die resultierende Artikel-Themen-Matrix nur 4'919 statt 4'922 Konferenzbeiträge beinhaltet. Die drei Konferenzbeiträge, die beim Topic Modeling verloren gegangen sind, weisen nur einen – meist zu kurzen - Titel als Textobjekt auf und konnten deshalb von der Text-Visualisierungs-Toolbox (TVT) in keines der 20 Themen eingeteilt werden. Es handelt sich dabei um «*Hyperchina: Adventures in hypermapping*» von Huffmann (1993), «*Columbus*» von Moellering & Ahlqvist (2013) und «*Strasbourg*» von Giraud (2013).

<sup>6</sup> <http://www.nltk.org/book/ch02.html> (Letzter Zugriff 21.04.2015)

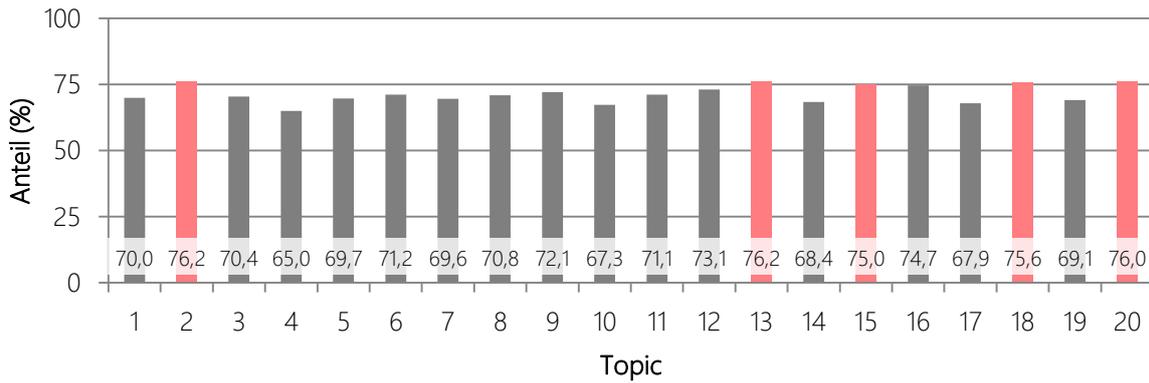


Abbildung 3-3: Relative Häufigkeit der einzelnen Topics in der gesamten Datengrundlage

Als Übersicht kann man zusätzlich die prozentuale Häufigkeit der Themen auf die Artikel ausrechnen, wie in Abbildung 3-3 gezeigt. Diese zeigt auf, in wie vielen Prozent der 4'919 Konferenzbeiträge das jeweilige Topic mit einem Wert in der Artikel-Themen-Matrix über Null und somit vertreten ist. Beim dem behandelten Beispiel (Tabelle 3-4) wären die Themen 1,2 und 4 mit 100% vertreten, während das Topic 3 in diesem Beitrag gar nicht (0%) vertreten ist.

Die y-Achse zeigt die prozentuale Häufigkeit des einzelnen Topics auf alle bearbeiteten Konferenzbeiträge. Auf diese Weise kann man erkennen, welche Themen in mehr Artikeln vertreten sind. Erkennbar ist auch die Problematik dieser Datengrundlage, in welcher jedes Topic in mehr als der Hälfte der Konferenzbeiträge repräsentiert wird. Es ist schwierig, das schon ziemlich fokussierte Datenfundament in klar abgegrenzte Themengebiete zu unterteilen. Mit einer breiter gestreuten Grundlage würde diese Unterteilung eindeutiger ausfallen (Tang et al. 2014). So weisen die Topics 2, 13, 15, 18 und 20 alle sogar einen Anteil von mehr als 75 Prozent auf. Das bedeutet, dass diese Topics mindestens von drei aus vier Konferenzbeiträgen repräsentiert werden. Wenn man zusätzlich die Tabelle 3-5 betrachtet, erkennt man, dass es sich um ziemlich breitgestreute Themen handelt. Diese Tabelle zeigt uns dadurch die Streuung der Themen auf die verschiedenen Konferenzbeiträge. Möchte man nun aber die durchschnittliche Bedeutung des einzelnen Themas auf alle Konferenzbeiträge hochgerechnet analysieren, so erkennt man, dass für die Topics 6, 17 und 20 die durchschnittliche Wahrscheinlichkeit auf alle Konferenzbeiträge am höchsten ist. In Kombination mit der Tabelle 3-5 erkennt man die Logik hinter dieser Verteilung. Die Wahrscheinlichkeit, dass die Begriffe dieser Themen in Kombination auftreten, ist am höchsten. Das wiederum bedeutet, dass es die am besten definierten Topics sind. In Anbetracht der Themeninhalte macht dies Sinn. Denn Themen wie *data*(6), *project*(17) oder *system*(20) können in verschiedenen Fachgebieten der kartographischen Wissenschaft eine bedeutende Rolle spielen (Abbildung 3-4).

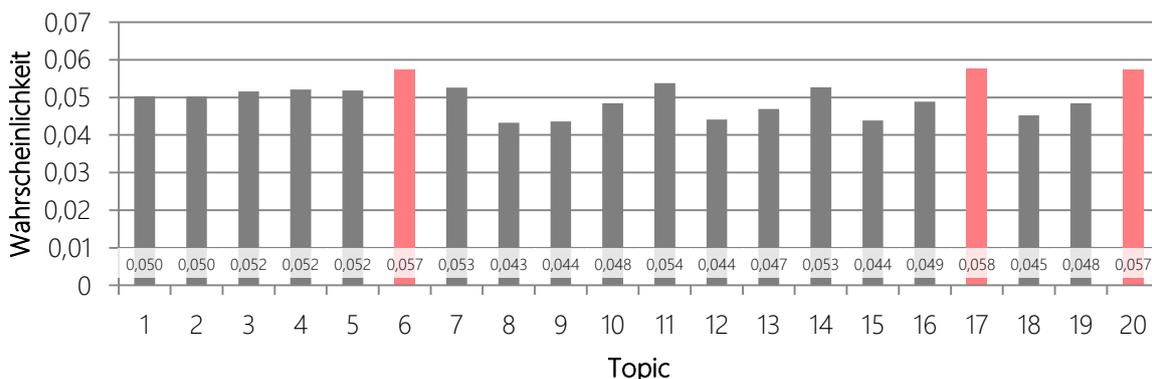


Abbildung 3-4: Durchschnittliche Wahrscheinlichkeit pro Topic

### 3.4 VISUALISIERUNGSPROZESS

Die im Kapitel 3.3 beschriebenen Abstracts und/oder Titel sollen nun analog zum ersten Gesetz der Geographie von Tobler (1970) in einer Informationslandschaft dargestellt werden. Das bedeutet, dass Artikel, die näher zueinander verwandt sind, was auf dieser Datengrundlage eine höhere semantische Ähnlichkeit bedeutet, in dem nicht-geographischen Informationsraum näher dargestellt werden. Neben der Distanz eignen sich auch weitere geographische Metaphern wie zum Beispiel die Region, um in diese Informationslandschaft übertragen zu werden. Die Methodik, um ein solches Resultat zu erreichen, wird in diesem Kapitel beschrieben. Es handelt sich dabei um ein wissenschaftlich etabliertes Vorgehen, welches in diesem Kapitel mit dem gesamten Prozess begründet und erklärt wird. Wie im Kapitel 2 gezeigt ist, befasst sich diese Arbeit nur mit der SOM-Technik als Dimensionsreduktions- und Gruppierungsmethode. Die resultierende Informationslandschaft wird dann für die Studie so bearbeitet, dass ein Vergleich einer diskretisierten und einer kontinuierlichen Visualisierung möglich ist. Wie man aus den Grundlagedaten eine durch geographische Metaphern erodierte Informationslandschaft erstellt, ist in diesem Kapitel beschrieben. In der Abbildung 3-5 ist ein Flussdiagramm des gesamten Visualisierungsprozesses zu sehen. Die Datenaufbereitung, welche im Flussdiagramm den grünen Sektor darstellt, wurde im Kapitel oben beschrieben. In den folgenden Unterkapiteln werden die einzelnen

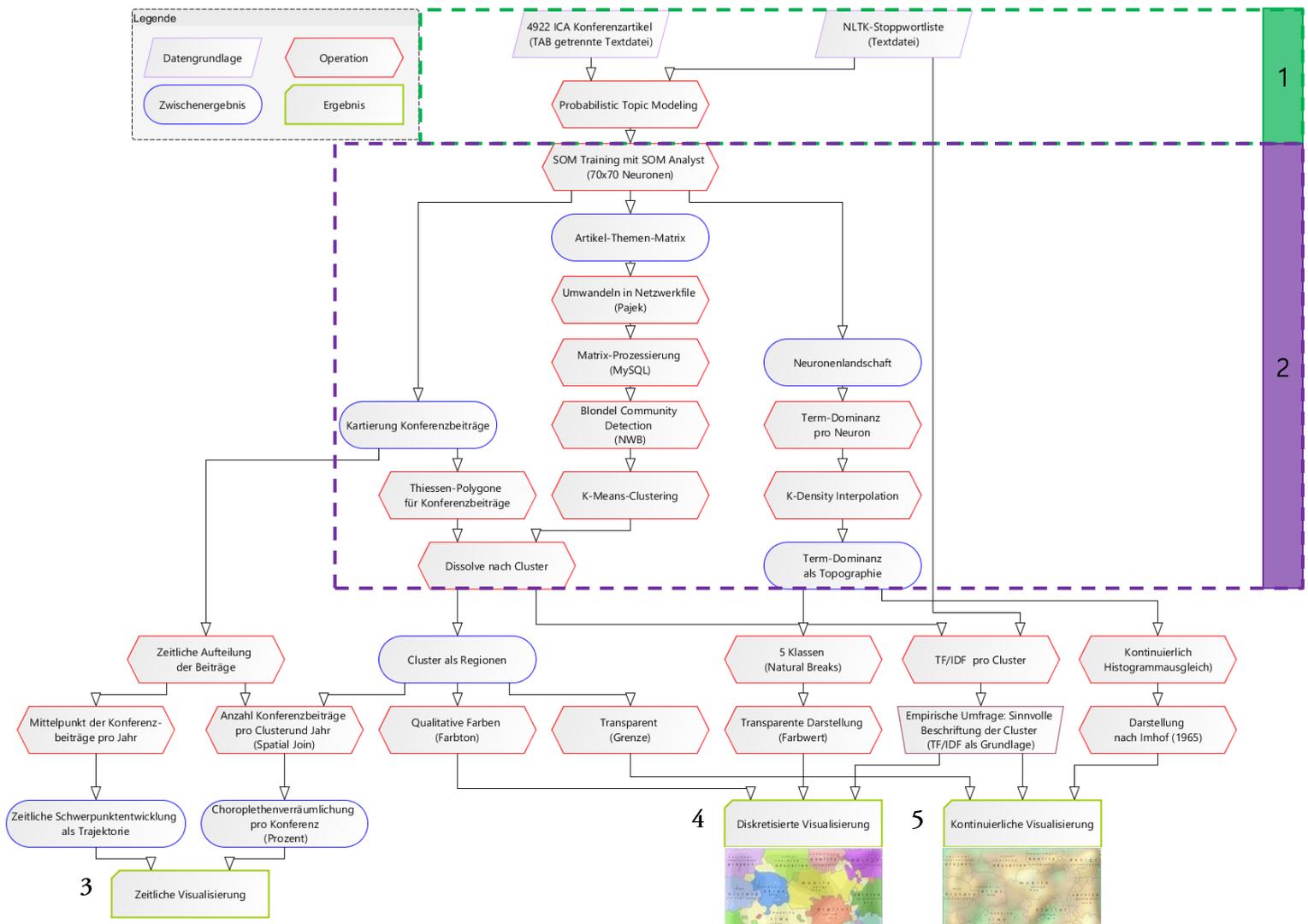


Abbildung 3-5: Flussdiagramm der verschiedenen Visualisierungsprozesse

Arbeitsschritte, die von der Datenaufbereitung bis zur finalen Visualisierung führen, detailliert erklärt.

Zuerst ist der generelle Vorgang beschrieben, welcher für das Erstellen einer Informationslandschaft notwendig ist. Bei Fragen in der Vorgehensweise gelten andere Forschungsarbeiten als Entscheidungsgrundlage (Fabrikant & Salvini 2011; Bruggmann 2012; Bruggmann et al. 2013a). Im Flussdiagramm (Abbildung 3-5) erkennt man die violette Markierung (2), welche die allgemeinen Prozessschritte beinhaltet. Die spezifischen Änderungen der visuellen Variablen, um zu den in der Studie verwendeten Visualisierungstypen zu gelangen, werden jeweils in den dafür vorgesehenen Kapiteln 3.4.2 & 3.4.3 beschrieben. Im Flussdiagramm werden mit den Ziffern 4 und 5 auf die beiden in der Studie verwendeten Visualisierungstypen hingewiesen. Die Ziffer 3 beschreibt zwei Möglichkeiten der Darstellung unter Berücksichtigung der zeitlichen Dimension.

### 3.4.1 ALLGEMEINER ARBEITSPROZESS

Die Inputdaten sollen nun mithilfe der resultierenden Artikel-Themen-Matrix des Topic Modeling in einen Informationsraum projiziert werden. Nach der Reduktion der Datendimension, die schon mit dem Topic Modeling begonnen hat, müssen die 20 Dimensionen (Themen) weiter reduziert werden, um die Daten auf eine zweidimensionale Karte zu projizieren. Hierfür wird die Methode der Self-Organizing-Maps (SOM) verwendet. Weitere angewandte Techniken und Methoden sind in diesem Kapitel analog des Flussdiagramms (Abbildung 3-5) beschrieben.

#### 3.4.1.1 Self-Organizing Maps (SOM)

Die aus dem Topic Modeling resultierende Artikel-zu-Thema-Matrix dient als Input der SOM-Berechnung mit der *SOM Analyst Toolbox* in *ArcGIS*, welche von Lacayo-Emery (2011) entwickelt wurde. Die verwendeten Parameter lassen sich von Lacayo-Emery (2011) und von Skupin & Esperbé (2011) ableiten. Einzelne Anpassungen für das Training und die Grösse der SOM werden vorgenommen um die Lesbarkeit der Daten zu optimieren. Wie in Skupin & Esperbé (2011) beschreiben, ist die Wahl der Anzahl Neuronen gleich bedeutend mit dem Clustering der Beiträge – wählt man eine höhere Anzahl an Neuronen, verteilen sich die Artikel mehr, wählt man eine kleinere Anzahl, so gruppieren sich die Beiträge mehr. Damit sich detaillierte Strukturen im Attributraum entwickeln können und die Neuronen nicht als Clustering-Ansatz genutzt werden, sollen alle Inputvektoren zumindest statistisch gesehen die Chance haben, ein Neuron für sich zu beanspruchen (Skupin & Esperbé 2011). Bei 4'922 Konferenzbeiträgen ergibt das eine SOM-Dimension von 70x70 hexagonale Neuronen. Im Gegensatz zu rechteckigen Neuronen weisen die hexagonalen nur sechs Nachbarn auf, was wiederum zu glatteren Resultaten führt als bei einem rechteckigen Neuronen-Gitter mit jeweils acht Nachbarn.<sup>7</sup> Generell werden überwiegend hexagonale Neuronen verwendet – so auch in allen Visualisierungsbeispielen, die im Kapitel Visualisierungsbeispiele und veröffentlichte Studien beschrieben sind.

Ausgehend von dieser Grösse des initialen SOM, basierend auf der zwei-dimensionalen Artikel-Themen-Matrix müssen beim Training des SOM einzelne Parameter aus der Anleitung des beschriebenen Tools angepasst werden. Nach einem Probe-Durchlauf mit den empfohlenen Parametern von Lacayo-Emery (2011) wurden einzelne Parameter optimiert. Schlussendlich verwendet man für das erste SOM-Training folgende Parameter: 49'000 für die *run length*, 70 für die *neighborhood size* und 0.04 für die *initial learning*

---

<sup>7</sup> <http://www.cis.hut.fi/somtoolbox/documentation/somalg.shtml> (Letzter Zugriff: 31.07.2015)

*rate*. Für das zweite Training wird die *run length* auf 490'000, die *neighborhood size* auf 8 und die *initial learning rate* auf 0.03 angepasst. Die *run length* wurde angepasst damit die Eingangsvektoren durch das längere Training mehr als einmal dargestellt werden können und so auch der Einfluss späterer dargestellten Vektoren nicht vernachlässigt wird. Pro *run* wird jeweils ein Eingangsvektor in ein Neuron projiziert (Skupin & Esperbé 2011). Die *neighborhood size* wird entsprechend der vorher begründeten *SOM*-Dimensionen angepasst – durch eine hohe *neighborhood size* im ersten Training bilden sich globale Strukturen heraus, welche sich durch eine kleinere *neighborhood size* verfeinern lassen. Die *learning rate*, welche das Gleichgewicht zwischen Überspezialisierung einzelner Bereiche (zu hohe *learning rate*) und diffusen Strukturen im *SOM* (zu tiefe *learning rate*) regelt, wird von der Anleitung von Lacayo-Emery (2011) übernommen. (Salvini 2012)

Als Resultat ergibt sich folgender *SOM-Output* mit den beschriebenen Parametern. Dieser besteht aus den zwei im Kapitel 2.1.3 beschriebenen Komponenten. Einerseits das hexagonale Neuronengitter, welches den gewünschten Dimensionen entspricht und andererseits die *BMU*, welche die Konferenzbeiträge als rote Punkte möglichst ihrer Artikel-Themen-Matrix entsprechend auf die Neuronen verteilen. Aus den grauen Neuronen im Hintergrund wird schlussendlich die Landschaft abgeleitet.

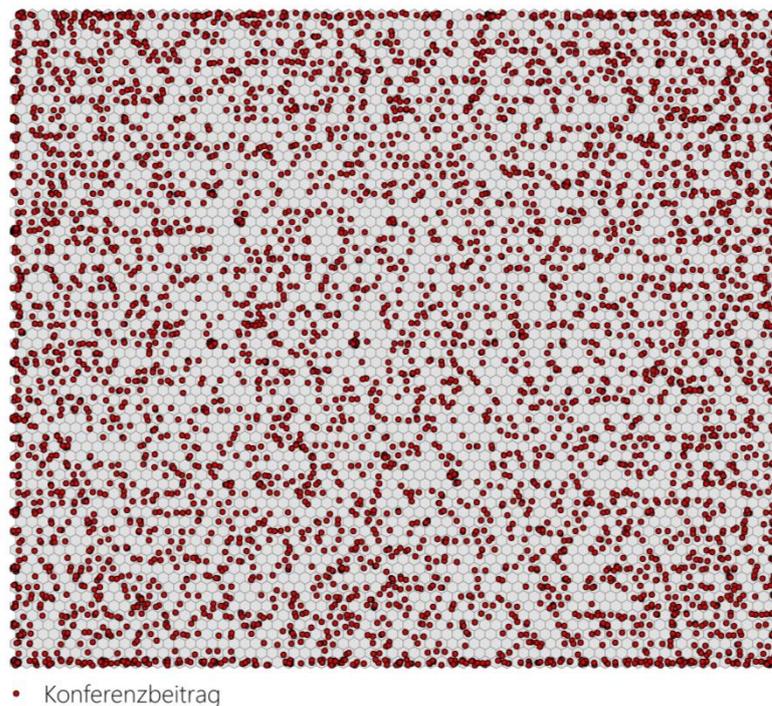


Abbildung 3-6: *SOM-Output*

### 3.4.1.2 Regionen in der Informationslandschaft

Die Region ist eine räumliche Taxonomie, welche den Raum basierend auf ähnlichen oder unterschiedlichen Eigenschaften eines Phänomens aufteilt (Fabrikant 2001). Diese Metapher führt zum Beispiel zu der politischen Gliederung der Welt. Diese besteht aus Kontinenten, die wiederum in Länder eingeteilt werden können. Dabei ist schon die hierarchische Struktur dieser Metapher ersichtlich, denn die Länder wiederum können in sich weiter aufgegliedert werden. So weisen Regionen jedem einzelnen Punkt im Informationsraum eine Bedeutung zu (Fabrikant 2001). Um die Informationen thematisch zu gliedern und die Grenzen der Regionen in der dazugehörigen Landschaft zu ziehen, müssen die Konferenzbeiträge mittels einer Clusteranalyse in Cluster zusammengefasst werden. Es handelt sich dabei um eine Aggregation als hierarchische Organisation der Daten (Fabrikant & Skupin 2005). Um eine sinnvolle Anzahl der Cluster aus der Natur der Daten zu begründen, wird der Blondel Community Algorithmus angewandt (Blondel et al. 2008). Dieser gruppiert die Konferenzbeiträge, wenn sie untereinander viele Beziehungen aufweisen und separiert sie, wenn sie kaum miteinander verbunden sind (Bruggmann, Salvini & Fabrikant 2013b). Diese Gruppierung basiert somit auf der semantischen Ähnlichkeit der Artikel. Durch diesen Algorithmus wird im Sinne des Modularitätsprinzip die Gruppierungsmethode so optimiert, dass die semantische Ähnlichkeit von Artikeln innerhalb der Gruppen möglichst hoch und zu Artikeln in anderen Gruppen möglichst klein ist (Blondel et al. 2008). Um die Beziehungen der Konferenzbeiträge untereinander zu vergleichen, wandelt man als Teil der geometrischen Generalisierung zunächst die bimodale Artikel-Themen-Matrix aus dem Topic Modeling in eine gewichtete Artikel-Ähnlichkeits-Matrix um. Die Tabelle 3-6 zeigt eine vereinfachte Form des Outputs vom Topic Modeling. Dabei stehen die Zeilen für die einzelnen Konferenzbeiträge und die Spalten für die verschiedenen Topics. Daraus folgend geben die Zellen an, wie stark ein Artikel zu einem Topic verbunden ist. Die Stärke aller thematischen Beziehungen eines Artikels beträgt aufsummiert immer 1. Aus diesem Grund ist in der Tabelle unten als Total immer 1 angegeben. Basierend auf diesen Relationen wird die Ähnlichkeit der Beiträge unter sich abgeleitet.

Tabelle 3-6: Vereinfachte exemplarische Artikel-Themen-Matrix II

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	...	Total
Artikel 1	0.4	0	0.1	0.5	0	...	1
Artikel 2	0.5	0	0	0.2	0.2	...	1
Artikel 3	0	0.2	0	0	0.5	...	1
...	...	...	...	...	...	...	1

Tabelle 3-7: Artikel-Ähnlichkeits-Matrix

	Artikel 1	Artikel 2	Artikel 3	...
Artikel 1	-	0.3	0	...
Artikel 2	0.3	-	0.1	...
Artikel 3	0	0.1	-	...
...	...	...	...	...

Es handelt sich dabei um einen mathematischen Prozess, bei welchem die thematischen Gewichtungen, die beide Artikel gemeinsam haben, multipliziert und über alle Themen aufsummiert werden. Im Beispiel (Tabelle 3-6) weisen die Artikel 1 und 2 bei Thema 1 und Thema 4 eine gemeinsame Gewichtung auf. Das führt dazu, dass die beiden Werte zu Thema 1 (Artikel 1: 0.4; Artikel 2: 0.5) multipliziert werden und zum Produkt derselben Berechnung für das Thema 4 (Artikel 1: 0.5; Artikel 2: 0.2) addiert werden. Es resultiert eine Ähnlichkeit von 0.3 als Summe von 0.2 (Thema 1) und 0.1 (Thema 4). Die Tabelle 3-7 zeigt die resultierende Artikelähnlichkeitsmatrix fundierend auf diesem Prinzip. Da Artikel 1 und 3 über

kein Thema verbunden sind, ergibt sich dabei keine Ähnlichkeit (0) zwischen diesen Artikeln. Zuletzt haben Artikel 2 und Artikel 3 eine Relation über das Thema 5 (Artikel 2: 0.2; Artikel 3: 0.5). Daraus resultiert eine Ähnlichkeit von 0.1 zwischen Artikel 2 und Artikel 3.

Der Arbeitsschritt beginnt mit dem Umwandeln der Matrix (Tabelle 3-6) in eine bimodale Pajek-Netzwerkdatei. Hierfür verwendet man *PAJEK* (Batagelj & Mrvar 2009), welches unter anderem dazu genutzt wird, um soziale Netzwerke zu analysieren. Die erstellte bimodale Netzwerkdatei benötigt man für das *Network Workbench Tool* (NWB-Team 2006). Es handelt sich dabei um eine frei verfügbare Software, die das Analysieren, Modellieren und Darstellen von Netzwerken ermöglicht. Ein Teil davon ist das Berechnen von Beziehungsstärke und – grad. Beide Attribute werden gewichtet aber ungerichtet in *NWB* (NWB-Team 2006) berechnet. Die resultierende Netzwerkdatei wird eine Textdatei umgewandelt, die wiederum in *MySQL* (2011) eingelesen und zu einer unimodalen Artikel-Ähnlichkeitsmatrix transformiert. Um spätere Berechnungen zu beschleunigen und auch die Clustererkennung zu vereinfachen, werden in diesem Schritt nur die 15 stärksten Beziehungen pro Artikel exportiert (Klavans & Boyack 2006). Dieser Prozess ist der semantischen Generalisierung unterzuordnen, in welcher die Grundgesamtheit der Relationen auf die bedeutendsten reduziert wird. Die resultierende Matrix wird nun via *txt2pajek*-Tool (Pfeffer 2013) in eine unimodale ungerichtete Netzwerkdatei (Pfeffer et al. 2013) umgewandelt und in *NWB* (NWB-Team 2006) auf Blondel Communities untersucht. Mit diesem Algorithmus bilden sich 18 Cluster – was in Anbetracht der *ICA*-Kommissionen durchaus Sinn macht. Wie bei Fabrikant & Salvini (2011) wird dieser Wert weiter verwendet, um die Anzahl Cluster für eine *k-Means*-Clusteranalyse in *SPSS* (IBM 2012) zu bestimmen und die Artikel eindeutig einem Cluster zuzuweisen. Über die Identifikationsnummer der einzelnen Artikel kann man diese Eigenschaft den Artikeln zuweisen. Um die Cluster zu differenzieren, werden die Konferenzbeiträge, welche in Abbildung 3-7 als Punkte dargestellt sind, bereits farblich unterschieden (Abbildung 3-7). Auf diese Weise kann man in einem ersten Schritt visuell überprüfen, ob aus der durchgeführten Clusteranalyse eine sinnvolle Ordnung resultiert. Diese Schritte werden bereits in *ArcMap* (ESRI 2014) durchgeführt. Wie man in Abbildung 3-7 erkennen kann, macht das Ergebnis durchaus Sinn.

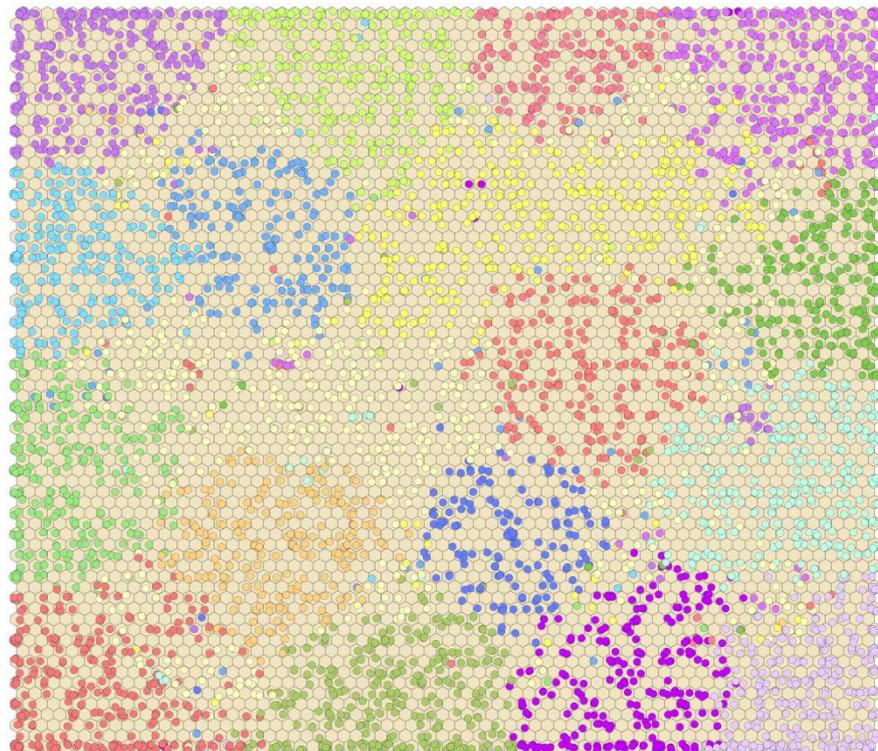


Abbildung 3-7: Resultat *K-Means* Clusteranalyse (n=18)

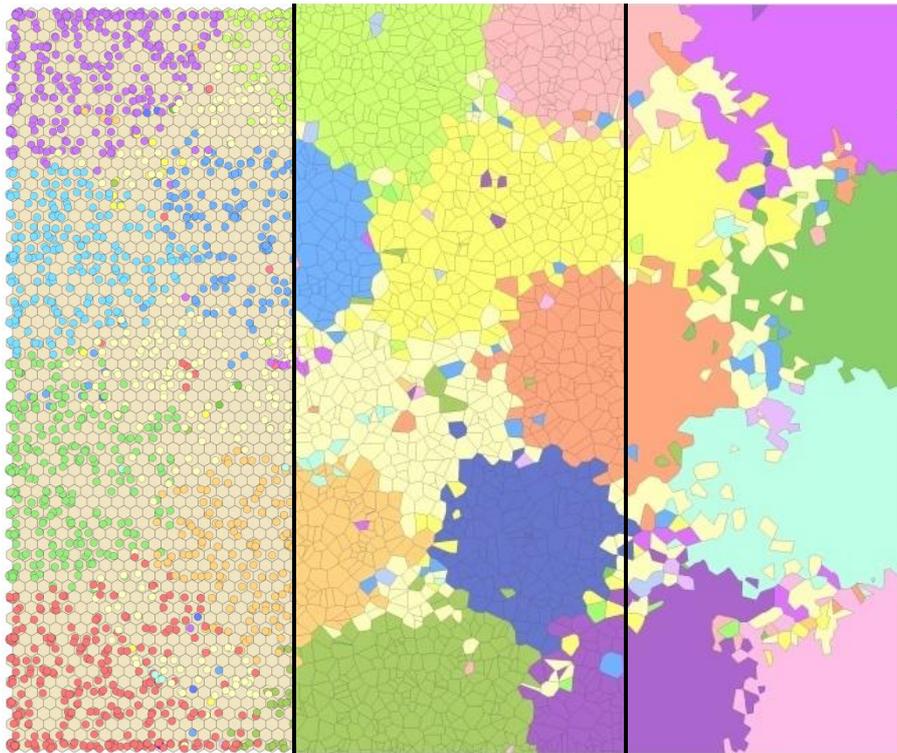


Abbildung 3-8: Arbeitsprozess von Punktdaten zu Regionen

Damit aus den Punktdaten der Konferenzbeiträge Regionen resultieren, muss man die komplette Fläche der Informationslandschaft jeweiligen Clustern zuweisen, damit nicht einzelne neutrale Gebiete entstehen, in welchen keine Konferenzbeiträge vorkommen. Als Lösung hierzu berechnet man die Thiessen-Polygone (oder auch Voronoi-Diagramme) der Konferenzbeiträge, welche in Abbildung 3-8 in der Mitte zu sehen sind. Dieses Vorgehen wurde auch in anderen Informationslandschaften angewandt (Skupin & de Jongh 2005; Skupin 2002a). Es handelt sich dabei um eine Form der geometrischen Generalisierung, in welcher die Punkte zu Polygonen zusammengefasst werden (Salvini 2012). Die qualitative Farbwahl der Cluster wird mit den unterschiedlichen Thematiken der *ICA* und der gewollten Anlehnung an eine politisch gegliederte Karte begründet.

Um die Lesbarkeit der Informationslandschaft zu verbessern, werden die Regionen analog kartographischer Techniken generalisiert. So werden einerseits Grenzen zwischen Thiessen-Polygone desselben Clusters mit dem *Dissolve*-Tool<sup>8</sup> in ArcMap (ESRI 2014) aufgelöst (Abbildung 3-8: rechts) und andererseits werden kleine thematische Exklaven mit dem *Eliminate*-Tool<sup>9</sup> aufgelöst. Wie man in der Abbildung 3-9 erkennen kann, werden auch die Grenzen der Polygone geglättet und vereinfacht. Dies kann man in ArcMap mit der *Bezd Simplify*-Methode<sup>10</sup> des «*Simplify Polygon*»-Tool Während die abgebildeten Polygone nicht generalisiert sind, zeigen die grauen Linien die generalisierten Grenzen und Regionen in der Informationslandschaft. Die Vereinfachung der geometrischen Formen als Teil der kartographischen Generalisierung reduziert die Komplexität der Informationslandschaft (Skupin 2000).

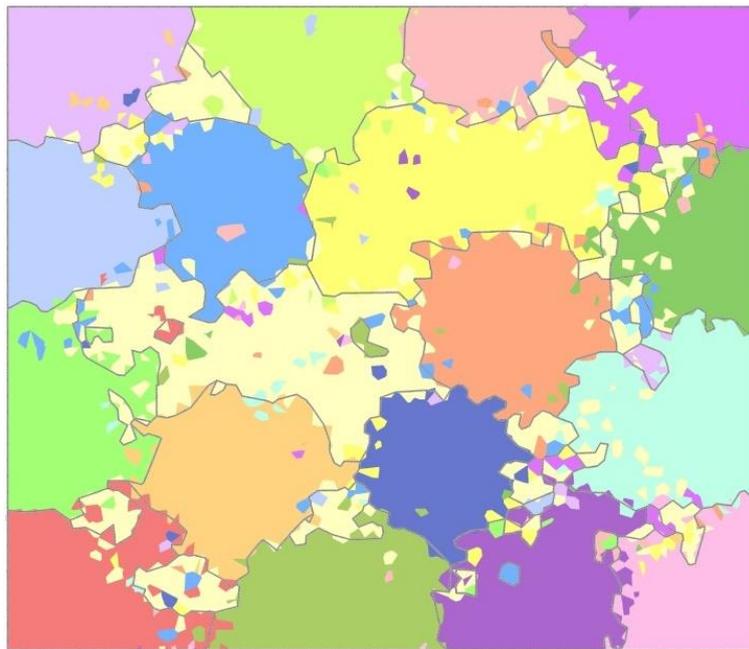


Abbildung 3-9: Quantitative Generalisierung

### 3.4.1.3 Beschriftung der Regionen in der Informationslandschaft

Wie bei gewöhnlichen Karten sollen die Regionen beschriftet werden, um die thematische Unterscheidung der Cluster hervorzuheben. Eine geeignete Beschriftung ist eine grosse Herausforderung in der Informationsvisualisierung. Die Frage nach der korrekten Platzierung der Labels und überhaupt nach einem charakteristischen Label ist schwierig zu beantworten (Skupin 2000). Für Letzteres besteht die Möglichkeit, repräsentative Begriffe pro Cluster aufgrund der Häufigkeit ihres Auftretens herauszufiltern und die Region mit diesen zu beschriften. Da die Datengrundlage thematisch fokussiert ist, stellt die Wahl dieser Begriffe eine grosse Herausforderung dar. Denn durch die thematisch homogene Datengrundlage sind einzelne Begriffe wie «*mapping*» oder «*data*» für mehrere Cluster häufig und können für eine sinnvolle Unterscheidung der Regionen nur begrenzt verwendet werden.

<sup>8</sup> <http://resources.arcgis.com/de/help/main/10.2/0017/00170000005n000000.htm> (Letzter Zugriff: 12.09.2015)

<sup>9</sup> <http://resources.arcgis.com/de/help/main/10.2/0017/00170000005p000000.htm> (Letzter Zugriff: 12.09.2015)

<sup>10</sup> <http://resources.arcgis.com/de/help/main/10.2/0070/00700000000110000000.htm> (Letzter Zugriff: 12.09.2015)

Um ein geeignetes Mass zu berücksichtigen, als nur die Auftretenshäufigkeit einzelner Begriffe, wird der *tf-idf* verwendet. Dieser verrechnet die *Term Frequency* mit der *Inverse Document Frequency*. Er indexiert also die Wichtigkeit einzelner Begriffe für ein Cluster im Verhältnis zu allen Begriffen in allen Cluster (Börner et al. 2003; Boyack et al. 2005). Da aber die Forschungsgebiete der *ICA* in sich geschlossen sind und 18 unterschiedliche Cluster beschriftet werden sollen, kann es auch mit dem *tf-idf*-Wert zu Überschneidungen kommen. Um dennoch eine sinnvolle Beschriftung der Regionen zu erreichen, wurden Personen aus dem universitärem Umfeld gebeten, die Repräsentativität einzelner Begriffe pro Cluster mithilfe des *tf-idf*-Werts zu beurteilen. Bei den ausgewählten Testpersonen dieser Vorstudie handelt es sich um eine Geographin und vier Geographen, die sich in GIS spezialisieren und sich somit mit den zu bearbeitenden Begriffen aber nicht unbedingt mit der dahinterstehenden Methodik auskennen. Ihnen wurde der Zweck dieser Vorstudie eine geeignete Beschriftung des jeweiligen Cluster abzuleiten sowie eine Erklärung der Unterlagen mündlich mitgeteilt. Bei den Unterlagen handelt es sich um 18 ausgedruckte Papiere, auf welchen jeweils Begriffe eines Clusters zu sehen sind.

Tabelle 3-8: Ausschnitt aus der Vorstudie

CLUSTER	TERM	TF-IDF	RANG	CLUSTER	TERM	TF-IDF	RANG
10	data	0.0030761	4	11	map	0.0055768	1
	national	0.002557	1		information	0.0014642	8
	information	0.0022073	3		cartographic	0.0010121	7
	spatial	0.0013222	4		data	0.0009851	16
	mapping	0.0011474	2		design	0.000883914	3

Dabei mussten die Testpersonen die Repräsentativität der fünf geeignetsten Begriffe pro Cluster zu bewerten. Als Vorlage dient ihnen eine Liste pro Cluster mit 20 Begriffen mit den höchsten *tf-idf*-Werten – sortiert nach diesem Wert von oben nach unten. In der Tabelle 3-8 sind als Beispiel zwei Cluster mit den fünf Begriffen mit den höchsten *tf-idf*-Werten abgebildet. Die Testpersonen mussten fünf der 20 Begriffe mit einer Punktzahl von eins bis fünf nach der jeweiligen Repräsentativität für das Cluster bewerten. Als Hilfe wurde noch die Rangierung des jeweiligen *tf-idf*-Wertes angegeben, welche den *tf-idf*-Wert des einzelnen Begriffes im Cluster mit dem Wert desselben Begriffes in den anderen Clustern vergleicht. Dies soll den Cluster übergreifenden Vergleich vereinfachen. Denn wie man im Beispiel (Tabelle 3-8) erkennen kann, hat das Wort *information* im zehnten den höheren *tf-idf*-Wert als im elften Cluster, obwohl der Begriff im letzteren auf einer höheren Stufe steht. Schlussendlich werden die Bewertungen summiert und die drei Begriffe mit der höchsten Punktzahl pro Cluster werden als Beschriftung verwendet.

Wie in Abbildung 3-10 zu sehen ist, wird der Begriff mit der höchsten Punktzahl durch einen fetten und grösseren Schriftzug hervorgehoben, während die beiden weiteren Begriffe das Forschungsgebiet des jeweiligen Clusters ergänzen. Die Platzierung, Ausrichtung und Formatierung der Beschriftung geschieht mit der *Maplex Label Engine*<sup>11</sup>, die in *ArcGIS* (ESRI 2014) zur Verfügung steht und die Regionen automatisiert beschriftet. Mit diesem Tool kann man jeweils verschiedene Layer mit unterschiedlichen Einstellungen beschriften. So wurde in dieser Arbeit zum Beispiel nur jeweils das grösste Polygon des Clusters angeschrieben und der Zeichenabstand an die Ausdehnung des jeweiligen Polygons angepasst. So sind die räumlichen Dimensionen für die Cluster nicht nur durch die Grenzen dazwischen sondern auch am Schriftzug erkennbar.



Abbildung 3-10: Beschriftung der Regionen

#### 3.4.1.4 Topographie der Informationslandschaft

Um eine sinnvolle Topographie der Informationslandschaft darzustellen, muss man zunächst die in den Daten liegenden Möglichkeiten abwägen. Als Optionen für diese Arbeit in Anbetracht der nachfolgenden Studie kamen zuerst eine interpolierte Landschaft der Anzahl Artikel in Frage. Begründen kann man diese Option mit dem Gedankenhintergrund, dass mehr Konferenzbeiträge auch einen höheren Stapel Papiere bedeuten würden und so die Relation zu einer natürlichen Landschaft mit Bergen und Tälern verständlich gemacht wird. Ein weiterer Lösungsansatz ist eine aussagekräftigere Variable, die man aus den vorhandenen Beiträgen berechnet. Die beschriebene Variable ist die *Begriffsdominanz*. Sie wird via Python-Script aus den Titeln und Abstracts der Konferenzbeiträge kalkuliert. Die Berechnung besteht aus dem Vergleich der Anzahl der häufigsten Wörter mit der Anzahl aller Wörter pro Neuron. Die Berechnungsgrundlage bilden die aggregierten Textbeiträge pro Neuron. Zusammengefasst zeigt die Variable die Wichtigkeit pro Neuron einiger häufiger Begriffe im Verhältnis zu allen Begriffen. Das

<sup>11</sup> <http://resources.arcgis.com/de/help/main/10.2/00s8/00s80000007m000000.htm> (Letzter Zugriff: 12.09.2015)

wiederum ist gleich bedeutend mit dem Fokus oder der Streubreite der Konferenzbeiträge. Es zeigt, ob sich die Forschungsartikel pro Neuron klar unterscheiden oder ob dieselben Begriffe von Bedeutung sind. So zeigt eine hohe Begriffsdominanz, dass wenige Begriffe pro Neuron sehr bedeutend sind. Dies weist auf einen thematischen Fokus in Daten hin. Haben die wichtigsten Begriffe keinen grossen Einfluss auf die Textgrundlage pro Neuron, ist die Begriffsdominanz tief. Auf diese Weise werden auch Neuronen, welche keine Konferenzbeiträge beinhalten, berücksichtigt. Sie bilden dann die Täler in dieser Informationslandschaft, da an diesen Stellen aufgrund der fehlenden Textgrundlage keine Begriffsdominanz vorhanden ist. Aufgrund bereits getätigter Arbeiten wird die Begriffsdominanz als Variable für die Oberfläche verwendet. (Skupin et al. 2013; Skupin & Fabrikant 2003; Skupin & de Jongh 2005; Skupin 2004, 2009)

Nach der Berechnung der Begriffsdominanz pro Neuron werden die berechneten Werte auf den Mittelpunkt des entsprechenden Neurons übertragen. Aus diesen Punktdaten und den damit verbundenen Begriffsdominanz-Werten wird die Oberfläche der Informationslandschaft durch eine Kerndichteanalyse der berechneten Begriffsdominanz in *ArcGIS* (ESRI 2014)<sup>12</sup> generiert. Das Resultat davon ist in Abbildung 3-11 zu sehen. Dabei erkennt man die helleren Bereiche, die eine höhere Begriffsdominanz übermitteln und die dunkleren Gebiete, welche die Täler dieser Landschaft darstellen sollen. Vergleicht man dazu die Verteilung der Konferenzbeiträge in Abbildung 3-6, dann macht diese Topographie durchaus Sinn.

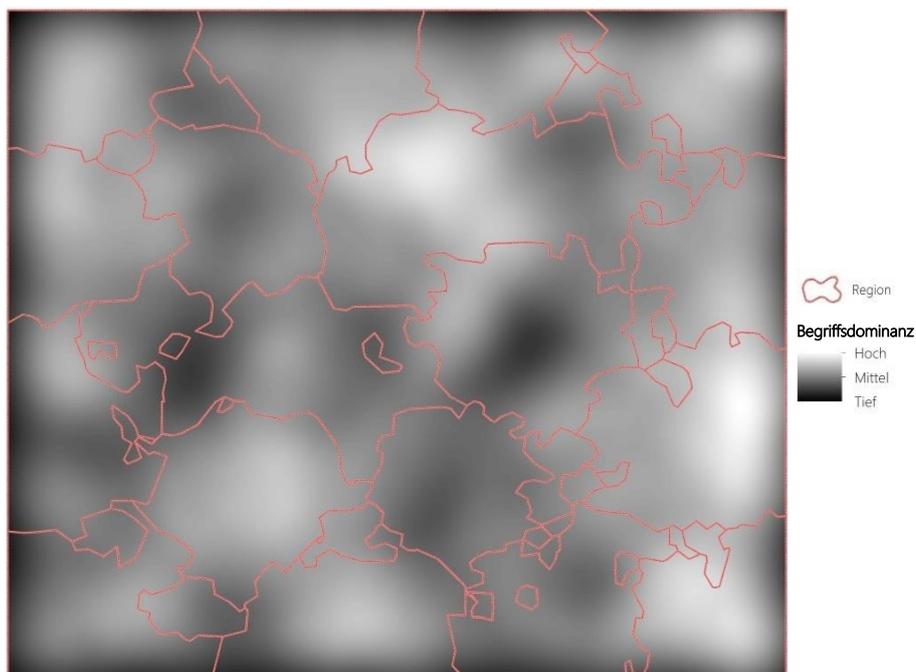


Abbildung 3-11: Begriffsdominanz-Landschaft

---

<sup>12</sup> <http://resources.arcgis.com/de/help/main/10.1/009z/009z0000000s000000.htm> (Letzter Zugriff: 23.07.2015)

### 3.4.2 VISUALISIERUNGSTYP D



Abbildung 3-12: Ausschnitt aus einer politisch-gegliederten, europäischen Karte

Da der Fokus in dieser Visualisierungsart wie bei einer gewöhnlichen politisch gegliederten Karte (Abbildung 3-12) mehr in der räumlichen Ordnung der Informationslandschaft als in der Oberflächeninformation liegt, werden die Regionen farblich unterschieden. Dennoch wird versucht, die Zusatzinformation der Begriffsdominanz in einer sinnvollen Form darzustellen. Als Lösung bietet sich eine klassierte und somit diskretisierte Ansicht der prozessierten, ursprünglich kontinuierlichen Daten an. (Tory & Moller 2004). So wird die Variable der Begriffsdominanz in fünf Klassen mit transparenten, sequentiellen Graustufen dargestellt (Brewer 1994). Bei der Klassierung wurde die *Natural Breaks*-Methode angewandt (Slocum et al. 2009). So kann man die Farbtonwahl für die Cluster qualitativ verschieden belassen und mittels des Farbwerts der einzelnen Regionen die Begriffsdominanz als einzelne Oberflächenniveaus darstellen (Fabrikant et al. 2006). Somit erscheint ein Farbton dunkler, wo dominante Wörter gemäss Begriffsdominanz in der Landschaft sind. Wenn man sich nun auf die visuellen Variablen von Bertin (1967) bezieht, wird für die qualitative Unterscheidung der Region der Farbton angewendet und für die Begriffsdominanz benötigt man den regionalen Farbton sowie den relativen Farbwert. Die Grenzen zwischen den Regionen werden durch den Visualisierungsvorschlag von politischen Grenzen in *ArcGIS* (ESRI 2014) dargestellt. Der gesamte hier beschriebene Visualisierungsprozess wird in *ArcGIS* (ESRI 2014) durchgeführt. Im weiteren Beschrieb dieser Arbeit wird die diskretisierte Visualisierung (Abbildung 3-13) als Visualisierungstyp D bezeichnet.



Abbildung 3-13: Finale diskretisierte Visualisierung (Visualisierungstyp D)

### 3.4.3 VISUALISIERUNGSTYP K



Abbildung 3-14: Ausschnitt aus einer europäischen Reliefkarte

Im Gegensatz zur diskretisierten Visualisierung steht bei dieser Darstellung die Kontinuität der Informationslandschaft im Zentrum. Vergleichen kann man diese Visualisierung mit einer gebräuchlichen Reliefkarte (Abbildung 3-14). Dabei erkennt man schon, dass die Gliederung in Regionen durch die Farbgebung an Bedeutung verliert. Die Farben für die kontinuierliche Begriffsdominanz entsprechen den empfohlenen Farben der Geländedarstellung von Imhof (1965). Damit dieselben Informationen wie bei der diskretisierten Informationslandschaft aufgezeigt werden, unterteilt man im Gegensatz zum Beispiel in Abbildung 3-13 die Landschaft durch ebenso visualisierte Grenzen. Damit die von Imhof (1965) gegebene Farbabfolge gleichmässig über das Informationsgelände verteilt ist, wird zur Visualisierung der Begriffsdominanz ein Histogrammausgleich vorgenommen. Zusätzlich fügt man einen Schummerungseffekt (*hillshade*)<sup>13</sup> ein, welcher die Betonung der Geländeoberfläche steigern soll. Kommt man nun auf die visuellen Variablen von Bertin (1967) zurück, ändert sich in dieser Visualisierungsart der Farbton für die Begriffsdominanz und der Farbwert für die Schummerung. Die Beschriftung und auch die Erscheinung der Grenzen werden für diese Visualisierung nicht geändert. In den weiteren Ausführungen zur Studie wird die kontinuierliche Darstellung (Abbildung 3-15) als Visualisierungstyp K bezeichnet.

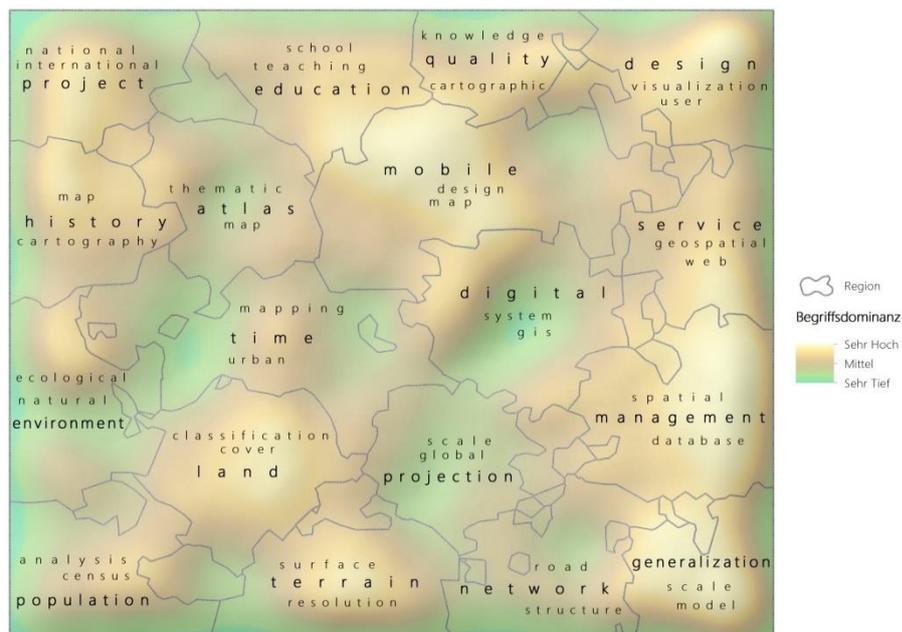


Abbildung 3-15: Finale kontinuierliche Visualisierung (Visualisierungstyp K)

<sup>13</sup> <http://resources.arcgis.com/de/help/main/10.1/009z/009z000000v0000000.htm> (Letzter Zugriff: 23.07.2015)

### 3.4.4 ZEITLICHE VISUALISIERUNG

Die Analyse der in der Datengrundlage verankerten zeitlichen Dimension kann verschieden gestaltet werden. Diese Arbeit enthält zwei Resultate, welche die Entwicklung des Inhalts in den Konferenzbeiträgen der *ICA* wiedergeben. Einerseits erkennt man in der Abbildung 3-16 die Verschiebung des Konferenzschwerpunkts entlang der zeitlichen Dimension. Diese Schwerpunkte werden mit *ArcGIS* (ESRI 2014) für jedes Jahr kalkuliert und mit einer gestrichelten Linie entlang der Zeitachse verbunden. Im Gegensatz zu Skupin (2009) und Skupin & Hagelman (2005) zeigt diese Darstellung nur die relative Verschiebung der Konferenzschwerpunkte und projiziert nicht die absolute Position der Konferenz in der virtuellen Informationslandschaft. In Zusammenhang mit der

Abbildung 3-17, welche die prozentuale Verteilung der Beiträge auf die resultierenden Cluster analog einer Choroplethenkarte darstellt, kann man die zeitliche Entwicklung der behandelten Thematiken nachvollziehen. So kann man erkennen, dass sich das Cluster «service|geospatial|web» erst um die Jahrtausendwende entwickelt hat. Auch interessant ist die einzige Fokussierung auf das umweltzentrale Thema «ecological|natural|environment» im Jahr 1997. Es ist auch das Jahr in dem das Kyoto-Protokoll<sup>14</sup> unterzeichnet wurde.

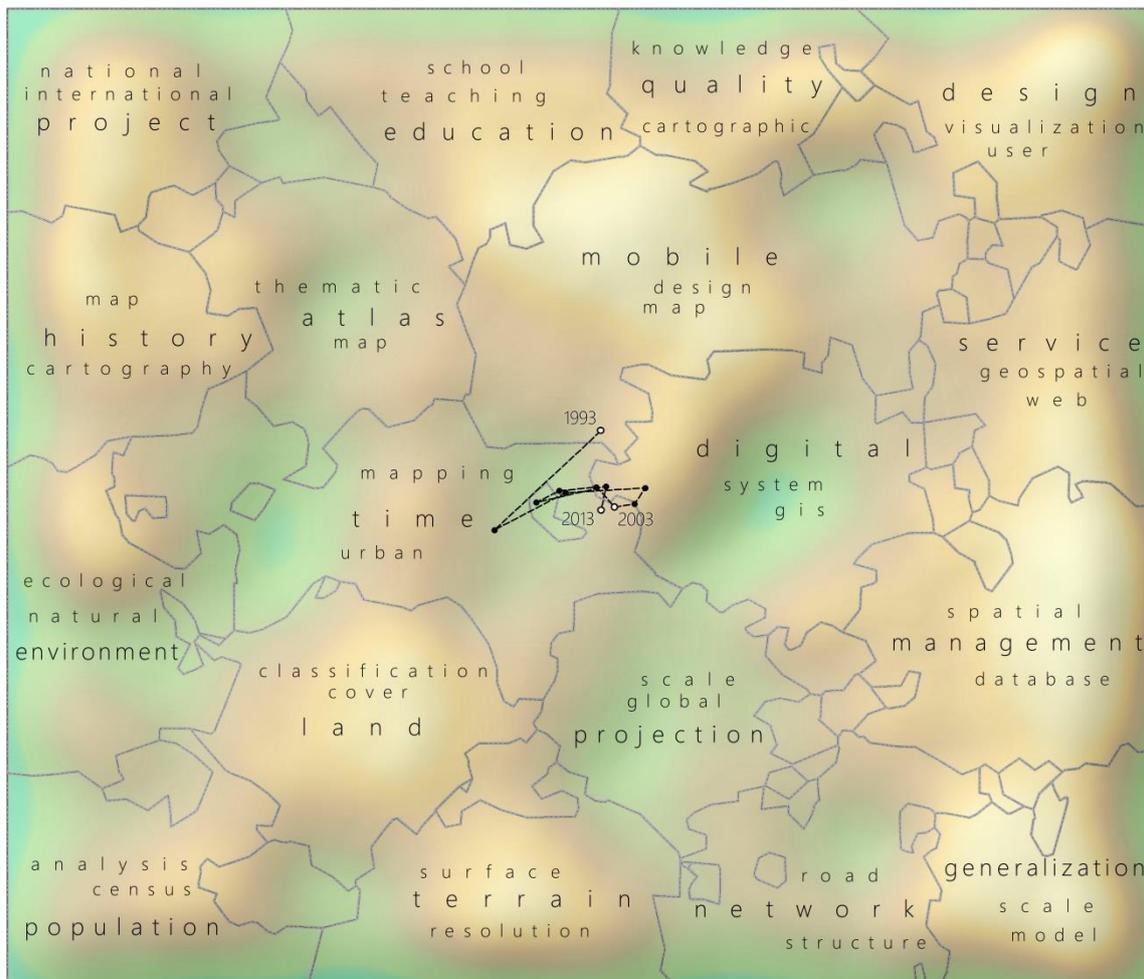


Abbildung 3-16: Schwerpunktverschiebung der *ICA*-Konferenzbeiträge von 1993-2013

<sup>14</sup> [http://unfccc.int/kyoto\\_protocol/items/2830.php](http://unfccc.int/kyoto_protocol/items/2830.php) (Letzter Zugriff: 09.09.2015)

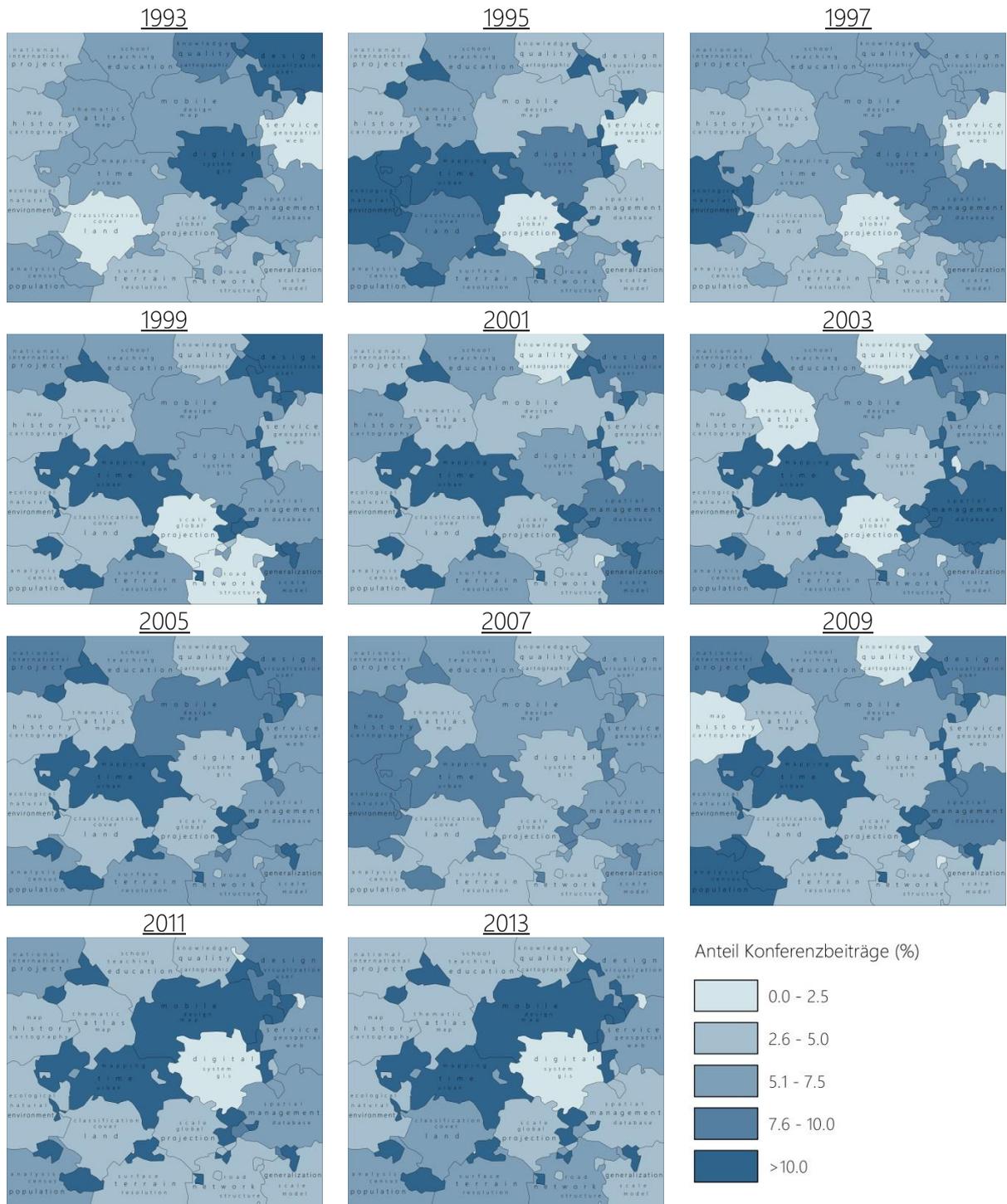


Abbildung 3-17: Verteilung der Konferenzbeiträge nach Jahr

## 3.5 STUDIENDESIGN

Um nun die auf solchen Darstellungen basierenden und aus den Fragestellungen resultierenden Hypothesen zu überprüfen, werden die beiden im Kapitel 3.4 erstellten Visualisierungstypen in einem Online-Experiment verglichen. Dieses Kapitel beschreibt die Stichprobe, die genutzte Plattform, verwendete Fragetypen sowie das gesamte Experimentelle Design.

### 3.5.1 STICHPROBE

Damit die Forschungsfragen beantwortet werden können, wird in dieser Arbeit eine Vielfalt an Probandinnen und Probanden benötigt. Aus diesem Grund wurden bestehende Kontakte genutzt und Personen aus dem universitären Umfeld sowie weitere anderweitig beschäftigte Kollegen und Kolleginnen per Mail eingeladen, an dieser Studie teilzunehmen. Schlussendlich haben 60 Personen an der Studie teilgenommen. Diese Stichprobengröße wird als genügend gross betrachtet. (Field 2013)

### 3.5.2 WEB-PLATTFORM

Wie schon beschrieben, konnte man an dieser Studie ohne Beaufsichtigung online teilnehmen. Da die Zeit dabei nicht berücksichtigt wird, kann die Betreuungsperson durch genügend ausführliche Anleitungen ersetzt werden. Dies wiederum steigert die Erreichbarkeit von möglichen Probandinnen und Probanden, da diese den Online-Fragebogen unabhängig von Zeit und Ort ausfüllen können. Aufgrund der Möglichkeiten im Aufbau, in der Auswertung und in der Darstellung wurde in dieser Arbeit die OnlineUmfragen.com-Plattform<sup>15</sup> verwendet. Der Fragebogen ist in drei Bereiche (Pretest, Hauptexperiment, Posttest) geteilt. Diese Bereiche werden jeweils mit Anleitungen und Erklärungen für die kommenden Aufgaben separiert. Weitere Beschreibungen zum Studienaufbau folgen in den nächsten Abschnitten.

### 3.5.3 STUDIENAUFBAU

Die Studie kann in verschiedene Bereiche eingeteilt werden, die wiederum mit kurzen Anleitungen eingeführt werden. Zuerst werden im Pretest persönliche Informationen abgefragt, um später sinnvolle Vergleiche zu ziehen. Im nächsten Bereich folgt das Hauptexperiment. Zum Schluss vergleicht man im Posttest neben der Präferenz mit Kontrollaufgaben die Antworten im Hauptteil mit einem räumlichen Ausschnitt der europäischen Landschaft.

#### 3.5.3.1 Pretest

Der Pretest verlangt demographische Informationen der Probandinnen und Probanden. Dazu zählen Eigenschaften wie Alter, Geschlecht, gewählte Studienrichtung, allfällige Sehschwächen oder der höchsten Bildungsgrad, der bisher erlangt wurde. Dazu wird die persönliche Einschätzung bezüglich Erfahrung mit Karten und Informationslandschaften auf einer Skala von 1 («schlecht») bis 5 («gut») abgefragt. Zum Abschluss werden im Pretest zwei Probeaufgaben gelöst, die wiederum zuerst erklärt sind. Bei den Probeaufgaben muss man aus einer Vorselektion den Konferenzbeitrag mit der höchsten oder niedrigsten Begriffsdominanz in der Informationslandschaft markieren. Da die Studie unbeaufsichtigt durchgeführt wird, müssen die Probandinnen und Probanden beide Probeaufgaben zunächst korrekt lösen, bevor sie mit dem Hauptexperiment beginnen können.

---

<sup>15</sup> OnlineUmfragen.com: <http://www.onlineumfragen.com> (Letzter Zugriff: 20.09.2015)

### 3.5.3.2 Hauptexperiment

Nachdem die Probeaufgaben richtig gelöst wurden, trennen sich die Fragebogen. Wie man in Abbildung 3-18 erkennen kann, ist der Ablauf der Fragetypen derselbe, nur der Visualisierungstyp unterscheidet sich. So erscheint bei Fragebogen A jeweils der Visualisierungstyp K zuerst, während bei Fragebogen B der Visualisierungstyp D als Erstes bearbeitet werden muss. Diese Abwechslung soll in der Abbildung 3-18 durch die unterschiedliche Abfolge mit dem Hintergrund in den Pfeilen «Übersicht» und «Ähnlichkeit» symbolisiert werden.



Abbildung 3-18: Hauptexperiment

Es gibt zwei Fragetypen im Hauptteil dieser Studie. Beim ersten Fragentyp handelt es sich um Übersichtsfragen, in welchen die Probandinnen und Probanden jeweils die Region mit den meisten oder wenigsten Konferenzbeiträgen aus einer vordefinierten Auswahl erkennen müssen. Zusätzlich muss man die Sicherheit bei der Beantwortung der Frage angeben. Ein solches Beispiel für beide Visualisierungstypen ist in Abbildung 3-19 zu sehen.

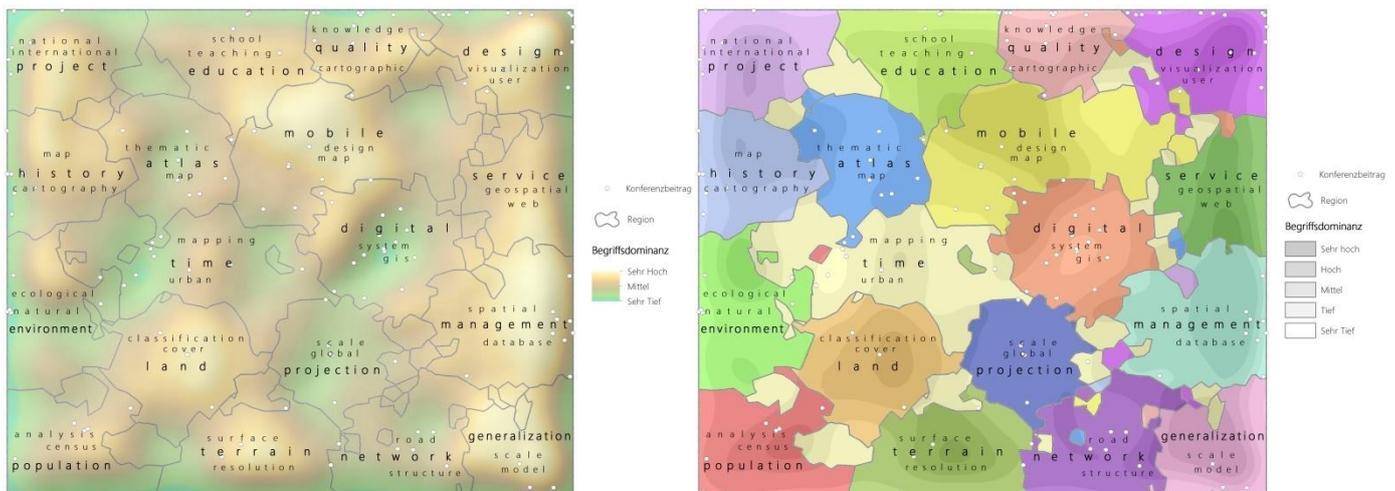


Abbildung 3-19: Beispielsvisualisierungen für Fragentyp «Übersicht» (links: K, rechts: D)

Dieser Fragentyp soll untersuchen, ob es einen Unterschied im überregionalen Punktmengenabgleich zwischen den Visualisierungstypen gibt. Es wird darauf geachtet, dass die Exklaven keine Rolle spielen in der Fragestellung, da die Zuordnung der Punkte in den Exklaven beim Visualisierungstyp K unmöglich ist. Trotzdem geht man davon aus, dass beim Visualisierungstyp D die Regionszuweisung für die Konferenzbeiträge, welche als weiße Punkte symbolisiert sind, besser funktioniert und sich die Testpersonen bei dieser Darstellung auch sicherer beim Beantworten der Fragen fühlen.

Da die relative Ähnlichkeit die Grundlage dieser Informationslandschaft bildet (Chalmers 1993), orientiert sich der zweite Fragentyp «Ähnlichkeit» an der Arbeit von Fabrikant et al. (2006). Die Testpersonen werden dabei gefragt, welcher Punkt aus A oder B dem Punkt 1 ähnlicher ist. Die dazugehörigen Visualisierungstypen finden sich in Abbildung 3-20. Die Antwort wird wie bei Fabrikant et al. (2006) auf einer Skala von 1 («Punkt A ist dem Punkt 1 viel ähnlicher») bis 9 («Punkt B ist dem Punkt 1 viel ähnlicher») angegeben. Zusätzlich werden die verschiedenen Einflüsse einzelner Faktoren wie Region, Distanz oder Oberflächenbeschaffenheit untersucht. Dabei müssen die Testpersonen den Einfluss des Faktors auf einer Skala von 1 («Keinen Einfluss») bis 4 («Viel Einfluss») einschätzen. Wie auch bei den Übersichtsaufgaben überprüft man die Sicherheit der Probandinnen und Probanden beim Beantworten der Aufgabe. Im Unterschied zu den Visualisierungen in Abbildung 3-20 haben die Testpersonen wie bei den Übersichtsaufgaben die gesamte Landschaft bearbeitet. Dieses Beispiel ist aus darstellungstechnischen Gründen zugeschnitten worden.

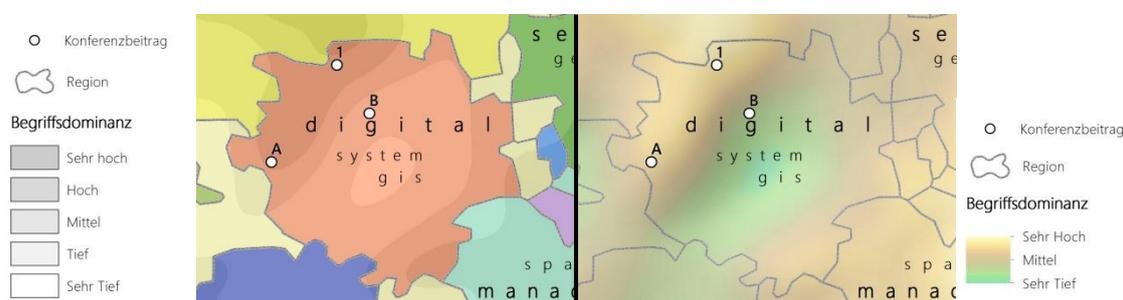


Abbildung 3-20: Beispielsvisualisierungen für Frage «Ähnlichkeit»

Dieser Fragentyp soll analysieren, ob bei einer der beiden Visualisierungstypen die Distanz als geographische Metapher einen grösseren Einfluss auf die Antwort hat. Dies kann man aus der Leistung der Testpersonen bei den Ähnlichkeitsaufgaben erkennen. Zusätzlich müssen die Einflüsse anderer Metaphern wie die Region selbst eingeschätzt und bewertet werden.

### 3.5.3.3 Posttest

Im Posttest wird in Kontrollaufgaben überprüft, ob die Probandinnen und Probanden die Ähnlichkeit von Punkten gleich bearbeiten bei räumlichen Daten. Hierbei dient ein Beispiel aus dem Schweizer Weltatlas (IKG ETH Zürich 2015). Dieses soll analog des Typs der Probeaufgaben und der Ähnlichkeitsaufgaben mit derselben Skala bearbeitet werden. Zusätzlich müssen die Testpersonen verschiedenen Strecken anhand der Distanz ordnen. Dieser Aufgabentyp überprüft den Einfluss der Topographie auf die Distanzwahrnehmung. Neben dem Abgleich zur europäischen Landschaft werden die Testpersonen gebeten, ihre persönlichen Präferenzen für die beiden Visualisierungstypen je nach Fragestellung darzulegen. Danach können die Probandinnen und Probanden noch freiwillig ihre Kontaktinformationen angeben, wenn sie über die Resultate der Studie informiert werden möchten. Ganz zum Schluss bedankt man sich herzlich für die Teilnahme.

### 3.5.4 EXPERIMENTELLES DESIGN

Da in den vorhergehenden Unterkapiteln der Studienaufbau mit allen Bereichen und Fragetypen bereits beschrieben und erklärt wurde, kann das experimentelle Design mit der effektiven Durchführung des Experiments ausgelegt werden.

#### 3.5.4.1 Versuchsdurchführung

Alle möglichen Kandidatinnen und Kandidaten aus dem privaten oder universitären Umfeld erhielten eine direkte Studieneinladung, die man nicht beantworten musste. Die Einteilung, welchen Fragebogen aus A oder B (Abbildung 3-18) die Teilnehmerinnen und Teilnehmer bekamen, basiert auf Zufall – und im späteren Verlauf auch auf bewusstem Vergeben, so dass beide Fragebogen von gleich vielen Testpersonen vollständig ausgefüllt wurden. Durchgeführt wurde die Studie im Zeitraum vom 6. bis 20. Juli 2015.

Anhand einer kurzen Erklärung der behandelten Thematik in der Studie sowie der ungefähren Zeitdauer konnten sich die Angeschriebenen für oder gegen eine Teilnahme entscheiden. Nach ihrem Einverständnis erklärt eine ausführliche Anleitung die generierte Informationslandschaft. Diese einheitliche Ausführung ist auch während dem Beantworten der Fragen jederzeit abrufbar. Die Fragen können ohne zusätzliche Hilfsmittel beantwortet werden. Falls trotzdem Fragen oder Probleme aufgetreten sind, konnte man mich über die angegebene E-Mail-Adresse oder Telefonnummer kontaktieren. Eine anonymisierte Auswertung wird von der Plattform OnlineUmfragen.com<sup>16</sup> mit der Zuweisung von Identifizierungsnummern garantiert. Nach der Zustimmung werden zunächst demographische Informationen wie das Alter oder Geschlecht gesammelt und persönliche Erfahrungen abgefragt, die einen Einfluss auf das Lösen der Aufgaben haben könnten. Nach einer Erklärung der Informationslandschaften und den beiden Probeaufgaben folgt der erste Teil des Hauptexperiments. Dieser wiederum ist unterteilt in 4 Fragen zur Übersicht. Die Visualisierungstypen sind nach dem im *KDDK*-Prinzip für Fragebogen A und *DKKD* für Fragebogen B angeordnet. Als Vorbild dient das *ABBA*-Prinzip aus der Arbeit von Martin (2008). Im zweiten Teil folgen sechs Fragen zur Ähnlichkeit, wobei jede Frage zu beiden Visualisierungstypen gestellt wird. Die Reihenfolge der insgesamt 12 Fragen ist zufällig, wobei darauf geachtet wird, dass dieselbe Frage zum anderen Visualisierungstyp in einem Abstand von mindestens zwei Fragen nicht beantwortet werden muss. Die abschliessende Reihenfolge der Fragen kann man aus der Tabelle 3-9 entnehmen.

Tabelle 3-9: Reihenfolge der Visualisierungstypen für beide Fragebogen

Fragebogen	Übersicht				Ähnlichkeit											
	Ü1	Ü2	Ü3	Ü4	Ä1	Ä3	Ä6	Ä5	Ä1	Ä4	Ä3	Ä5	Ä2	Ä4	Ä6	Ä2
A	K	D	D	K	D	K	K	K	K	D	D	D	K	K	D	D
B	D	K	K	D	K	D	D	D	D	K	K	K	D	D	K	K
<i>K</i>	<i>Visualisierungstyp K</i>															
<i>D</i>	<i>Visualisierungstyp D</i>															

Die Kontrollaufgaben nach dem Experiment stellen die Verbindung zu den räumlichen Landschaften dar. Sind diese beantwortet, können die Testpersonen ihre vom Aufgabentyp abhängige Präferenz für einen Visualisierungstyp angeben. Die Umfrage endet mit der Danksagung.

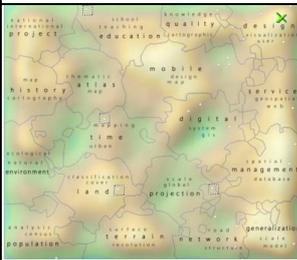
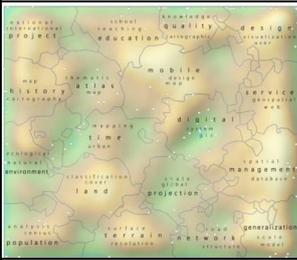
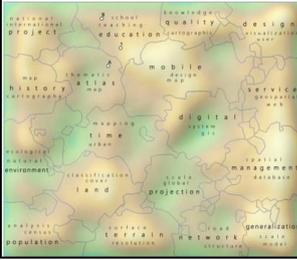
<sup>16</sup> OnlineUmfragen.com: <http://www.onlineumfragen.com> (Letzter Zugriff: 20.09.2015)

In der gesamten Auswertung wurde der Faktor «Zeit» vernachlässigt, da einerseits ohne Aufsicht nicht überprüft werden kann, ob die Testperson ihre Zeit ausschliesslich der Studie widmet oder ob weitere Tätigkeiten die Beantwortung der Fragen verzögert haben und so die Auswertung verfälschen würde. Andererseits sollte der Vergleich zwischen den Visualisierungstypen hauptsächlich in der Betrachtungsebene und nicht in der Zeitebene stattfinden. Aus diesem Grund ist die durchschnittliche Zeit von 38 Minuten, welche die Probandinnen und Probanden zum Komplettieren der Studie im Schnitt benötigten, nicht aussagekräftig.

#### 3.5.4.2 Unabhängige Variable

Alle Teilnehmerinnen und Teilnehmer sehen beide Arten der Informationslandschaft. Es handelt sich also um ein *Within-Subject*-Studien-Design (Martin 2008). Die unabhängige Variable ist also die Darstellungsart der Informationslandschaft. Die eine ist in einer diskretisierten Variante (Visualisierungstyp D) und die andere als kontinuierliche Landschaft (Visualisierungstyp K) visualisiert. Wie die beiden inhaltlich gleichen Typen aufgebaut sind und inwiefern sie sich unterscheiden wird im Kapitel 3.4 beschrieben. Wie man in der Tabelle 3-9 erkennen kann, sehen die Testpersonen für die Übersichtsfragen nur einen der beiden Typen, während der andere Fragebogen den jeweils anderen Typ enthält. Dies wird mit der Befürchtung eines Lerneffekts und mit dem Ziel einer rein prozentualen Richtig/Falsch-Auswertung begründet. Bei den Ähnlichkeitsfragen sehen die Testpersonen beide Visualisierungstypen zur selben Frage, allerdings in einer zufälligen Reihenfolge und mit mindestens zwei weiteren Fragen zwischen derselben Frage. Die Studie wurde mithilfe vom Buch «*Doing Psychology Experiments*» von David W. Martin (2008) gestaltet. Eine Übersicht von Beispielsvisualisierungen zu den einzelnen Aufgabentypen sowie die dazugehörigen abhängigen Variablen ist in der Tabelle 3-10 dargelegt.

Tabelle 3-10: Experimentelles Design

		Unabhängige Variable	
Abhängige Variablen		Visualisierungstyp D	Visualisierungstyp K
Aufgabenstellung	Probefragen		
	Übersichtsfragen		
	Ähnlichkeit		

### 3.5.4.3 Abhängige Variablen

#### 3.5.4.3.1 Leistung

Die Leistung der Testpersonen in den einzelnen Tasks wird verschieden gemessen. In den Probeaufgaben geht es darum, die Frage richtig zu beantworten – also gibt es bei dieser Aufgabenstellung das klassische Richtig/Falsch-Schema. Da es sich bei diesen Fragestellungen nur um Probeaufgaben handelt, werden diese in der Auswertung nicht berücksichtigt. Dasselbe Prinzip gilt aber auch bei den Übersichtsaufgaben. Dabei kann aufgezeichnet werden, welcher Visualisierungstyp generell eher zu einer korrekten Antwort führt. In der Tabelle 3-10 ist deshalb bei beiden Aufgabentypen die abhängige Variable Korrektheit beschrieben. Bei den Ähnlichkeitsfragen müssen die Probandinnen und Probanden wie erwähnt die Ähnlichkeit dreier Punkte vergleichen. Dieser Wert bezeichnet die Wahrnehmung der Ähnlichkeit und wird entsprechend ausgewertet. Diese Variable wird in der Tabelle 3-10 als Performance angegeben. Bei diesem Task ist der nähere Punkt auch immer ähnlicher und somit die korrekte Antwort. Zusätzlich bietet die Einschätzung verschiedener Faktoren (Regionszugehörigkeit, die Regionsgrenzen, die Begriffsdominanz, die Distanz zwischen den Konferenzbeiträgen und die Oberflächenbeschaffenheit der Landschaft) eine detailliertere Auswertungsmöglichkeit. Mit dieser Auswahl müssen die Bedeutung verschiedener geographischer Metaphern wie die Region mit ihren Grenzen oder auch die Distanz oder der Weg zwischen ausgewählten Punkten selber eingeschätzt werden.

### 3.5.4.3.2 Sicherheit

Bei beiden Aufgabentypen wird nach der Performance auch jeweils gefragt, wie sicher die Testpersonen bei der Beantwortung der Fragen sind. Diese Variable überprüft, ob bei einem der beiden Visualisierungstypen das Lösen der Aufgabe vereinfacht oder ob sie vielleicht auch eine falsche Sicherheit übermittelt. Es handelt sich dabei wieder um ein Mass, welches die Probandinnen und Probanden selber einschätzen müssen.

### 3.5.4.3.3 Präferenz

Wie bereits erwähnt wird im Posttest gefragt, welche Darstellung einer Informationslandschaft für die einzelnen Fragetypen präferiert wird. Interessant ist auch die allgemeine Favorisierung von einem der zwei Visualisierungstypen.

### 3.5.4.4 Kontrollvariable

Die zu vergleichenden Stichproben beziehen sich in dieser Arbeit nur auf die Ausbildung. Denn wie schon Fabrikant et al. (2006) herausgefunden haben, gibt es bei der Ähnlichkeitswahrnehmung keinen Unterschied im Geschlecht. Und da die durchschnittliche Erfahrung in Bereichen wie Karten lesen oder gestalten mit dem Geographie-Studium positiv korreliert, vergleicht man nur den Einfluss des Geographie-Studiums (Abbildung 3-21). Der Korrelationskoeffizient beträgt  $-0.454$  und ist auf einem Signifikanzniveau von  $0.001$  zweiseitig signifikant.

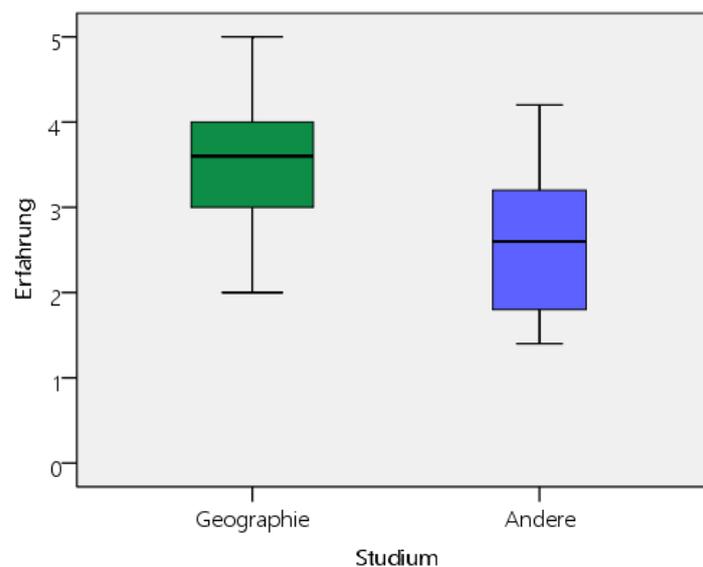


Abbildung 3-21: Boxplot zur durchschnittlichen Erfahrung abhängig von der Ausbildung

### 3.5.4.4.1 Geographie-Studium

Eine der Fragestellungen dieser Arbeit ist unter anderem der Vergleich zwischen Geographie-Studentinnen und Studenten und Testpersonen ohne eine geographische Ausbildung. Aus diesem Grund sind im Pretest Fragen zur beruflichen Ausbildung enthalten. Das Verständnis von geographischen Metaphern dürfte den Geographinnen und Geographen wegen der beinahe täglichen Auseinandersetzung mit solchen leichter fallen. Dabei stellte sich die Frage, ob Nebenfach-Geographie-Studentinnen und Studenten nun zu der Stichprobe mit einer geographischen Ausbildung gehören oder nicht. Damit die Anzahl Probandinnen und Probanden für beide Gruppen ausgeglichen ist, werden die Nebenfach-Studentinnen und Studenten deshalb in die andere Gruppe eingeteilt.

### 3.5.5 STATISTIK

Die Resultate der Studie sind statistisch signifikant ab einem  $p$ -Wert von 0.05. Durchgeführt wurde die statistische Analyse mit SPSS (IBM 2012). Da man neben dem Visualisierungstyp als unabhängige Variable auch verschiedene Kontrollvariablen nutzt, wird von Field (2013) eine mehrfaktorielle gemischte Varianzanalyse (*Mixed-design ANOVA*) vorgeschlagen. So können die verschiedenen Faktoren zwischen den Visualisierungstypen und einer Kontrollvariablen verglichen und analysiert werden. Während die deskriptive Statistik jeweils das arithmetische Mittel und einen positiven und negativen Standardfehler enthält und in einem Diagramm dargestellt ist, wird die gemischte Varianzanalyse mit der  $F$ -ratio ( $F$ ), dem Freiheitsgrad ( $df$ ), den  $p$ -Werten und der Stärke der Beziehung ( $\eta_p^2$ ) beschrieben. Für die Gestaltung der Diagramme wurde Microsoft Excel (Microsoft 2010) verwendet und für den Boxplot benötigte man das Statistikprogramm SPSS (IBM 2012). Die Skalen für die einzelnen Fragetypen sind in der Tabelle 3-11 abgebildet.

Tabelle 3-11: Skalenübersicht

Fragetyp		Auswertung	
		Minimum	Maximum
Übersichtsfragen	Auswahl	0 («Falsch»)	1 («Richtig»)
	Sicherheit	1 («Sehr unsicher»)	6 («Sehr sicher»)
Ähnlichkeitsfragen	Ähnlichkeit	1 («Punkt 1 und A sind sich viel ähnlicher»)	9 («Punkt 1 und B sind sich viel ähnlicher»)
	Einflüsse	1 («Keinen Einfluss»)	4 («Viel Einfluss»)
	Sicherheit	1 («Sehr unsicher»)	6 («Sehr sicher»)
Kontrollfragen	Auswahl	0 («Falsch»)	1 («Richtig»)
	Ähnlichkeit	1 («Punkt 1 und A sind sich viel ähnlicher»)	9 («Punkt 1 und B sind sich viel ähnlicher»)
	Reihenfolge	1 («Längste Distanz»)	4 («Kürzeste Distanz»)
Präferenz	Visualisierungstyp	1 («Ziehe Visualisierungstyp D am meisten vor»)	5 («Ziehe Visualisierungstyp K am meisten vor»)

## 4 RESULTATE

Dieses Kapitel zeigt die Ergebnisse der durchgeführten Studie. Die beiden untersuchten Visualisierungstypen sind als Resultat am Ende des Kapitels 3.4.2 bzw. 3.4.3 zu sehen und werden hier nur in Zusammenhang mit einzelnen Fragetypen verwendet. Die Resultate werden in einzelnen Sektionen präsentiert. Dabei beginnen statistische Analysen bei der Stichprobenzusammensetzung. Danach geht man über zur Auswertung der grundlegenden Fragestellung – ob es einen Unterschied zwischen den beiden Visualisierungstypen gibt. Zum Schluss vergleicht man noch den Einfluss des Geographie-Studiums in der Wahrnehmung. Die beiden offenen Fragestellungen der Präferenz und der Sicherheit werden innerhalb der beiden Sektionen ausgewertet.

### 4.1 STICHPROBE

Wie schon beschrieben haben 60 Testpersonen an der Studie teilgenommen und dabei mehr als 90% der Studie ausgefüllt. Aufteilen lässt sich die Stichprobe mit einem durchschnittlichen Alter von 25,9 Jahren wie in der Abbildung 4-2 dargestellt. Dabei sind verschiedene Ausbildungen und Hintergründe der Testpersonen beabsichtigt. Die Spalte «Abschluss» zeigt den höchsten erlangten Abschluss, die die Testperson bisher erlangt hat. Bei «Ausbildung» lag der Schwerpunkt beim Unterschied vom Geographie-Studium und einer nicht-geographischen Ausbildung. Unter der Kategorie «Andere» gehören neben anderen Studiengängen wie Rechtswissenschaften auch andere Ausbildungen wie eine Lehrausbildung dazu. Die Verschiedenheit dieser Kategorie ist in Abbildung 4-1 zu sehen. Neben der demographischen Aufteilung der Stichprobe mussten die Teilnehmerinnen und Teilnehmer ihre persönliche Erfahrung in Themengebieten einschätzen, die das Lösen der gestellten Aufgaben beeinflussen könnten. Diese Aufteilung ist in Abbildung 4-4 zu sehen.



Abbildung 4-1: Die Aufteilung der Klasse «Andere» in der Kategorie «Ausbildung» (kreiert mit Wordle: [www.wordle.net](http://www.wordle.net))

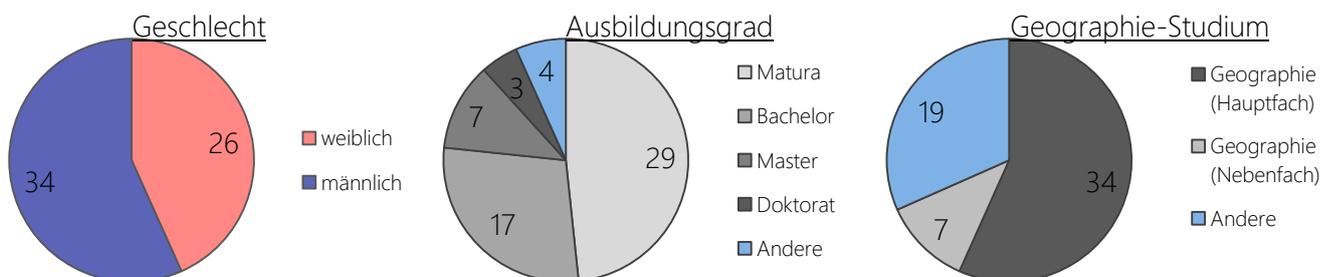


Abbildung 4-2: Aufteilung der Stichprobe nach Geschlecht, bisher höchster erlangter Bildungsgrad sowie nach geographischem Schwerpunkt

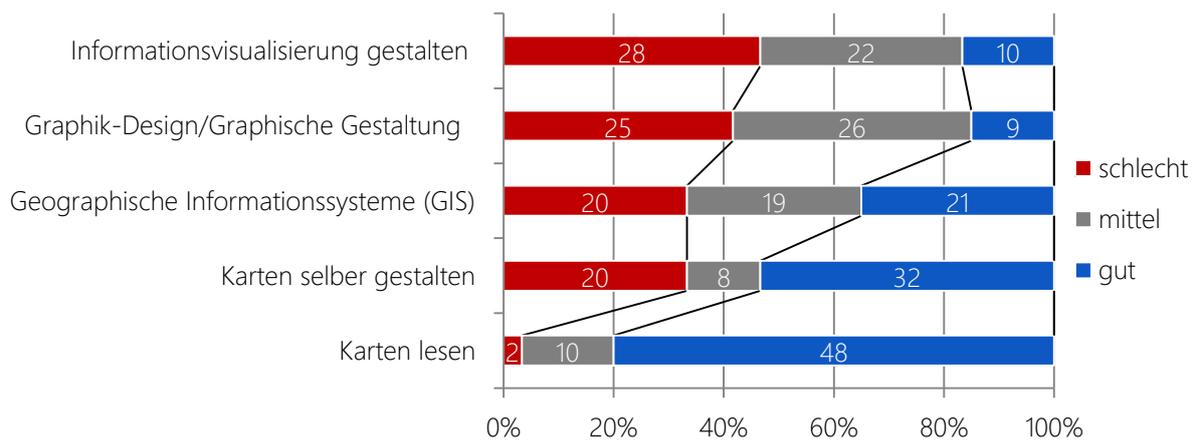


Abbildung 4-4: Aufteilung der Stichprobe nach eingeschätzten Kenntnissen [n=60]

Wie schon im Kapitel 3.5.4.4.1 beschrieben, beruht der in der Fragestellung beschriebene Vergleich des Einflusses einer geographischen Ausbildung zwischen einem (geplanten) Abschluss – also nur die Hauptfach- Absolventinnen und Absolventen – und einem sonstigen Abschluss. Die Studentinnen und Studenten, welche das Nebenfach Geographie belegen, werden bei der Auswertung zu den anderen Ausbildungen dazugezählt. Weiter können die 34 Hauptfach-Geographinnen und Geographen in ihre bereits abgeschlossenen oder geplanten Spezialisierungen unterteilt werden. Die Abbildung 4-3 zeigt die relativen und absoluten Verhältnisse für die Studentinnen und Studenten im Hauptfach Geographie.

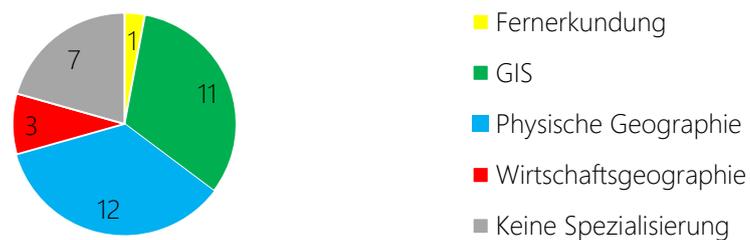


Abbildung 4-3: Aufteilung der geographischen Spezialisierungen des Hauptfachstudiums [n=34]

## 4.2 VISUALISIERUNGSTYP

Dieses Kapitel vergleicht nur die beiden Visualisierungstypen anhand der Performance und der Angaben der Testpersonen und berücksichtigt keine weiteren Faktoren. Dabei analysiert man je nach Fragestellung die Antworten und vergleicht die Resultate. Aus diesem Grund beginnt man dem Studienablauf entsprechend mit dem Unterschied bei den Übersichtsfragen und geht dann über zu den Ähnlichkeitsfragen.

### 4.2.1 ÜBERSICHT

Bei den Übersichtsaufgaben, in welchen die Punktmengen in den Regionen abgeglichen werden müssen, erkennt man aus dem Diagramm in Abbildung 4-5, dass es einen knappen Unterschied gibt zwischen den Visualisierungstypen. Bei den vier Fragen, wobei je zwei zu jedem Visualisierungstyp gestellt wurden, ergibt sich folgende Abbildung 4-5. Die Y-Achse stellt der Mittelwert richtig beantworteter Fragen dar, wobei das Maximum jeweils bei zwei liegt. Daraus lässt sich erkennen, dass im Schnitt die Fragen mit der diskreten Visualisierung besser beantwortet wurden – umgerechnet wurden 74% der Übersichtsfragen bei der kontinuierlichen und 77% bei der diskretisierten Darstellung korrekt beantwortet. Der Unterschied ist nicht signifikant [ $F(1,58)=0.534$ ,  $p>0.05$ ;  $\eta_p^2=0.009$ ].

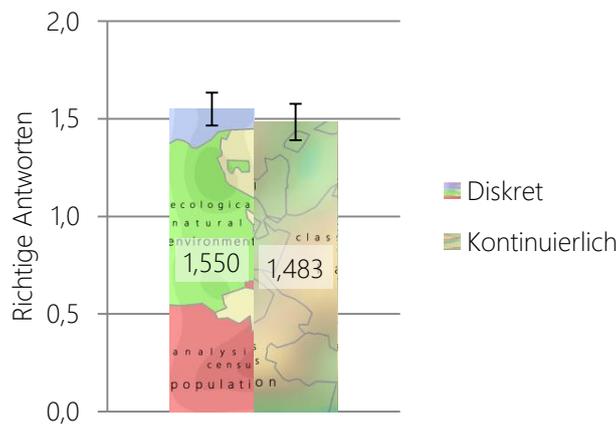


Abbildung 4-5: Vergleich der korrekten Antworten für die Übersichtsaufgaben [n=60] - Durchschnitt und  $\pm SF$

Neben dem korrekten Lösen der Aufgaben mussten die Teilnehmerinnen und Teilnehmer auch ihr Sicherheitsgefühl bei der Beantwortung der Fragen selbst einschätzen. Dabei mussten sie das Vertrauen in

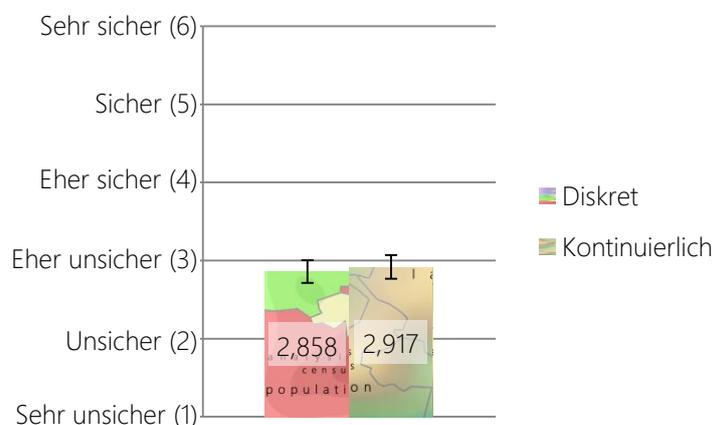


Abbildung 4-6: Vergleich der Sicherheit beim Beantworten der Übersichtsaufgaben [n=60] - Durchschnitt und  $\pm SF$

ihre Antwort auf einer Skala von 1 («Sehr unsicher») bis 6 («Sehr sicher») einordnen. Der Unterschied zwischen den Visualisierungstypen ist wiederum nicht signifikant [ $F(1,58)=0.509$ ,  $p>0.05$ ;  $\eta_p^2=0.009$ ]. Wie man in Abbildung 4-6 sieht, gibt es wiederum eine knappe Differenz. Allerdings widerspricht dieses Resultat mit der Anzahl richtiger Antworten – bei der kontinuierlichen Visualisierung wurden also im Schnitt nicht nur weniger Fragen richtig beantwortet, die Testpersonen fühlten sich dabei auch sicherer.

#### 4.2.2 ÄHNLICHKEIT

Im Gegensatz zu den Übersichtsaufgaben wurden bei der Ähnlichkeit die Testpersonen zusätzlich gefragt, wie stark der Einfluss verschiedener Faktoren auf ihre Antwort war. Zu diesen Faktoren gehörten die Regionszugehörigkeit, die Regionsgrenzen, die Begriffsdominanz, die Distanz zwischen den Konferenzbeiträgen und die Oberflächenbeschaffenheit der Landschaft. Zuerst wird das Resultat der Ähnlichkeitswahrnehmung auf der Skala von 1 («Punkt 1 und A sind sich viel ähnlicher») bis 9 («Punkt 1 und B sind sich viel ähnlicher») verglichen. Wobei der Punkt B immer der Punkt ist, welcher von Punkt 1 weiter entfernt ist. Vergleicht man zuerst allgemein nur die Ähnlichkeitsbewertung, ist der Unterschied zwischen den Visualisierungstypen mit einer asymptotischen Signifikanz von 0.077 knapp nicht bedeutend [ $F(1,58)=3.251$ ,  $p>0.05$ ;  $\eta_p^2=0.053$ ].

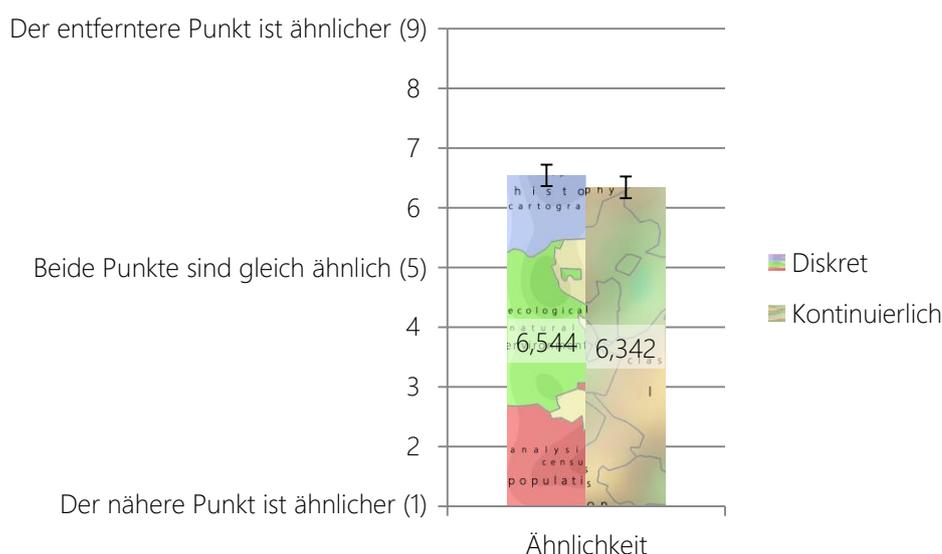


Abbildung 4-7: Vergleich der Ähnlichkeit beim Beantworten der Ähnlichkeitssaufgaben [n=60] - Durchschnitt und ±SF

In der Abbildung 4-7 ist der Unterschied ersichtlich, dass bei der diskreten Visualisierung die räumliche Distanz zwischen den Punkten eine noch kleinere Rolle spielt als bei der kontinuierlichen. Mit diesem Ergebnis vergleicht man nun zusätzlich die Einflüsse der einzelnen Faktoren, welche auch die Distanz beinhalten. Dabei wird zuerst der Schnitt über alle Faktoren verglichen um zu überprüfen ob überhaupt ein Unterschied zwischen den Visualisierungstypen vorhanden ist. In der Analyse ist erkennbar, dass mit dem Durchschnittswert aller Faktoren kein signifikanter Unterschied resultiert [ $F(1,58)=1.260$ ,  $p>0.05$ ;  $\eta_p^2=0.021$ ]. Grundsätzlich ist die Informationslandschaft anhand der Distanz als relatives Ähnlichkeitsmass strukturiert, was bedeutet, dass die Fragen mehrheitlich falsch beantwortet wurden. Beim Visualisierungstyp K wurden trotzdem im Schnitt 19,1% und beim Visualisierungstyp D 18,1% der jeweils 6 Fragen in diesem Sinne korrekt beantwortet. Dabei muss man sagen, dass alle Werte unter 5 als korrekt und alle darüber als falsch eingestuft werden [ $F(1,58)=0.165$ ,  $p>0.05$ ;  $\eta_p^2=0.003$ ].

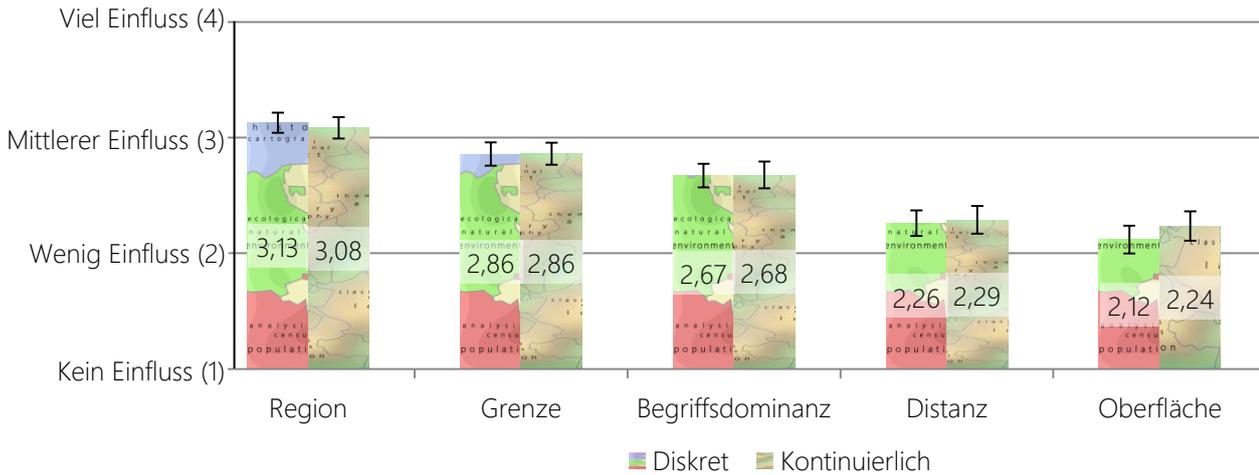


Abbildung 4-8: Vergleich des Einfluss der einzelnen Faktoren [n=60] - Durchschnitt und ±SF

Trotz fehlender Signifikanz vergleicht man in einem weiteren Schritt den Einfluss der einzelnen Faktoren (Abbildung 4-8). Die Skala dafür reicht dabei von 1 («Keinen Einfluss») bis 4 («Viel Einfluss»). Wie man schon in der Abbildung 4-8 erkennt, gibt es kaum Unterschiede bei dem Einfluss der Grenze, Begriffsdominanz und der Distanz. Die Regionszugehörigkeit wird bei der diskreten Visualisierung als ein wenig einflussreicher eingeschätzt als bei der kontinuierlichen Darstellung. Der auffälligste Unterschied ist allerdings bei der Oberfläche zu erkennen – diese wird, wie man in Tabelle 4-1 erkennen kann, mit signifikant mehr Einfluss eingeschätzt.

Tabelle 4-1: Statistische Auswertung der einzelnen Faktoren

Faktor	F(1,58)	p	$\eta_p^2$
Regionszugehörigkeit	F(1,58)=0.453	p>0.05	$\eta_p^2=0.008$
Regionsgrenzen	F(1,58)=0.000	p>0.05	$\eta_p^2=0.000$
Begriffsdominanz	F(1,58)=0.016	p>0.05	$\eta_p^2=0.000$
Distanz	F(1,58)=0.236	p>0.05	$\eta_p^2=0.005$
Oberflächenbeschaffenheit	F(1,58)=4.190	p<0.05	$\eta_p^2=0.067$

Wie bei den Übersichtsaufgaben mussten die Testpersonen ihre Sicherheit auf einer Skala von 1 («Sehr unsicher») bis 6 («Sehr sicher») einschätzen. Bei den Ähnlichkeitsaufgaben ist das Resultat in Abbildung 4-9 dargestellt. Bei diesem Fragentyp fühlen sich die Testpersonen im Durchschnitt sicherer bei der kontinuierlichen Visualisierung als bei der diskreten Variante. Dieser Unterschied ist allerdings nicht signifikant [F(1,58)=1.159, p>0.05;  $\eta_p^2=0.020$ ].

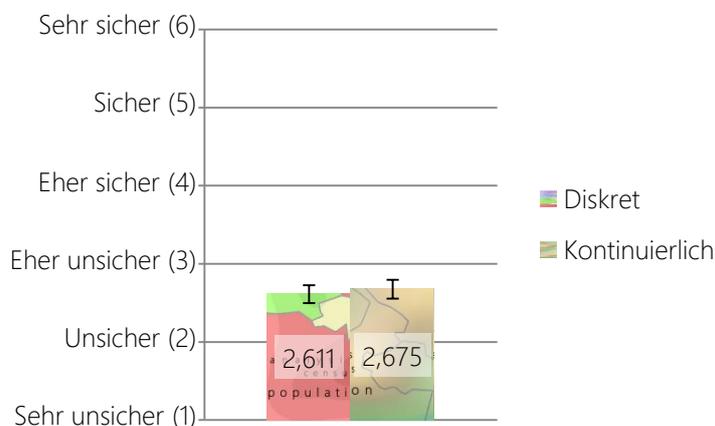
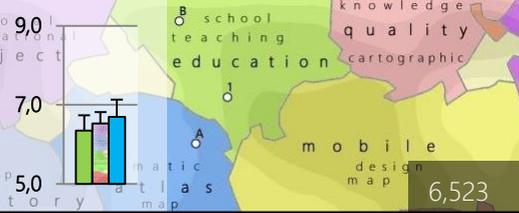
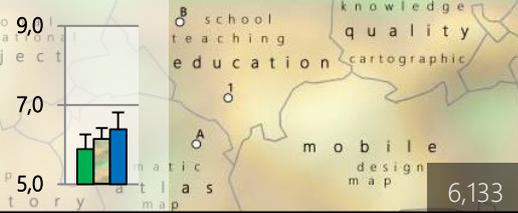
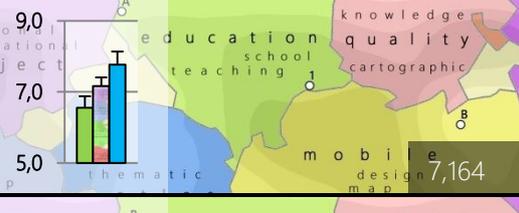
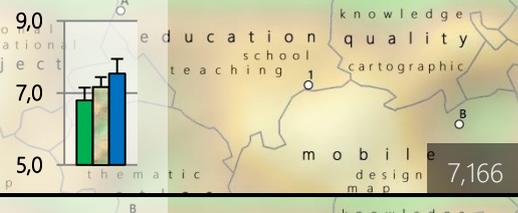
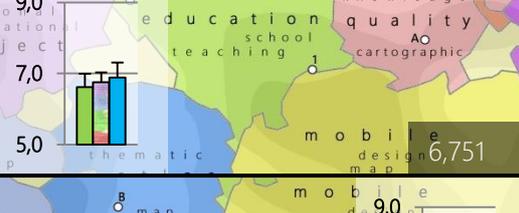
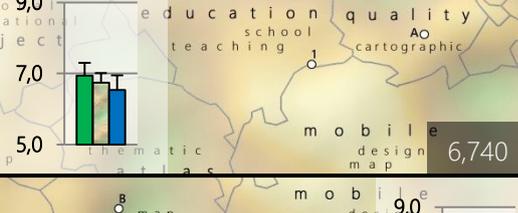


Abbildung 4-9: Vergleich der Sicherheit beim Beantworten der Ähnlichkeitsaufgaben [n=60] - Durchschnitt und ±SF

### 4.2.3 AUFGABENÜBERSICHT (ÄHNLICHKEIT)

Neben dem durchschnittlichen Unterschied bietet der Aufgabentyp der Ähnlichkeitswahrnehmung eine weitere Möglichkeit an, die Darstellungen zu vergleichen. Dabei wird die Ähnlichkeit signifikant abhängig vom Aufgabentyp bewertet [ $F(5,290)=11.933$ ,  $p<0.05$ ;  $\eta_p^2=0.171$ ]. Aus diesem Grund sind die Unterschiede je nach Aufgabentyp und Ähnlichkeitsbewertung in der Tabelle 4-2 aufgeführt. Dabei kann man auch schon die Tendenzen erkennen, die im Kapitel 4.2.2 bereits ausgewertet wurden. Wie zum Beispiel, dass praktisch durchgehend bei der diskreten Visualisierung die räumliche Distanz für die Ähnlichkeitswahrnehmung eine kleinere Rolle spielt als bei der kontinuierlichen. Um zu schauen, ob die Ähnlichkeitswahrnehmung signifikant vom Visualisierungstyp beeinflusst wird, vergleicht man für jede Aufgabe die Mittelwerte der Antworten in einer mehrfaktoriellen gemischten Varianzanalyse (*Mixed-design ANOVA*). Die Spalte der Statistik in der beschriebenen Tabelle dient zur statistischen Auswertung. Daneben zeigt die Auswertung graphisch der Gegensatz zwischen den Geographinnen und Geographen (Grüner Balken) und den Nicht-Geographinnen und Nicht-Geographen (Blauer Balken) für die jeweilige Visualisierung. Der Balken zwischen den beiden Gruppen stellt den Mittelwert der gesamten Stichprobe dar, welcher im schwarzen Rechteck jeweils auch exakt angegeben ist. Für jeden Balken ist der positive Standardfehler dargestellt. Zu beachten ist die Skala im Diagramm, welche immer die Ähnlichkeit zwischen dem Punkt 1 und dem weiter entfernten Punkt aus A und B aufzeigt. Welcher Punkt dies ist, steht in der ersten Spalte der Tabelle. Zusätzlich weist der positive Standardfehler im Diagramm auf die jeweilige Streuung in den Antworten hin. Ausserdem ist im Diagramm der kontinuierlichen Visualisierung von Aufgabe 4 der Balken für die Nicht-Geographinnen und Nicht-Geographen nicht dargestellt – der Wert für diesen Balken beträgt exakt 5 und entspricht deshalb der X-Achse. Der Schnittpunkt der beiden Achsen wurde nicht angepasst, da die Darstellungen sonst nicht mehr einheitlich und vergleichbar wären.

Tabelle 4-2: Übersicht zu der Beantwortung der einzelnen Fragen abhängig vom Visualisierungstyp [n=60]

Aufg.	Diskret	Kontinuierlich	Statistik
1 (9=B)			p>0.05 F(1,58)=2.054 $\eta^2=0.034$
2 (9=A)			p>0.05 F(1,58)=0.000 $\eta^2=0.000$
3 (9=B)			p>0.05 F(1,58)=0.002 $\eta^2=0.000$
4 (9=A)			p>0.05 F(1,58)=1.272 $\eta^2=0.021$
5 (9=A)			p<0.05 F(1,58)=6.001 $\eta^2=0.094$
6 (9=B)			p>0.05 F(1,58)=0.074 $\eta^2=0.001$

Der einzig signifikante Unterschied zwischen Visualisierungstyp K und Visualisierungstyp D besteht bei der fünften Aufgabe, bei welcher alle Punkte in derselben Region liegen. Dabei spielt die Distanz eine bedeutend stärkere Rolle bei der kontinuierlichen Darstellung. Die Begriffsdominanz dagegen scheint bei der diskreten Visualisierung die Ähnlichkeitswahrnehmung zu beeinflussen. Vergleicht man für die Aufgabe 5 noch die einzelnen Einflüsse, zeigt sich allerdings, dass dabei kein Faktor als signifikant einflussreicher bewertet wurde (Tabelle 4-3).

Tabelle 4-3: Vergleich der einzelnen Faktoren für Aufgabe 5

Regionszugehörigkeit	F(1,58)=1.333	p>0.05	$\eta_p^2=0.022$
Regionsgrenzen	F(1,58)=0.785	p>0.05	$\eta_p^2=0.013$
Begriffsdominanz	F(1,58)=0.427	p>0.05	$\eta_p^2=0.007$
Distanz	F(1,58)=0.000	p>0.05	$\eta_p^2=0.000$
Oberflächenbeschaffenheit	F(1,58)=0.094	p>0.05	$\eta_p^2=0.002$



### 4.2.5 PRÄFERENZ

Wie in den Fragestellungen schon beschrieben ist, möchte man die Präferenz der Probandinnen und Probanden für eine der beiden Visualisierungstypen überprüfen. Hierbei mussten die Testpersonen jeweils ihre Präferenz auf einer Skala von 1 («Ziehe Visualisierungstyp D am meisten vor») bis 5 («Ziehe Visualisierungstyp K am meisten vor») angeben. Aus diesem Grund schneidet bei der Abbildung 4-11 die vertikale die horizontale Achse bei 3,0. Ist der Durchschnittswert kleiner, wird die diskrete Visualisierung bevorzugt, ist er grösser, wird eher die kontinuierliche Visualisierung präferiert.

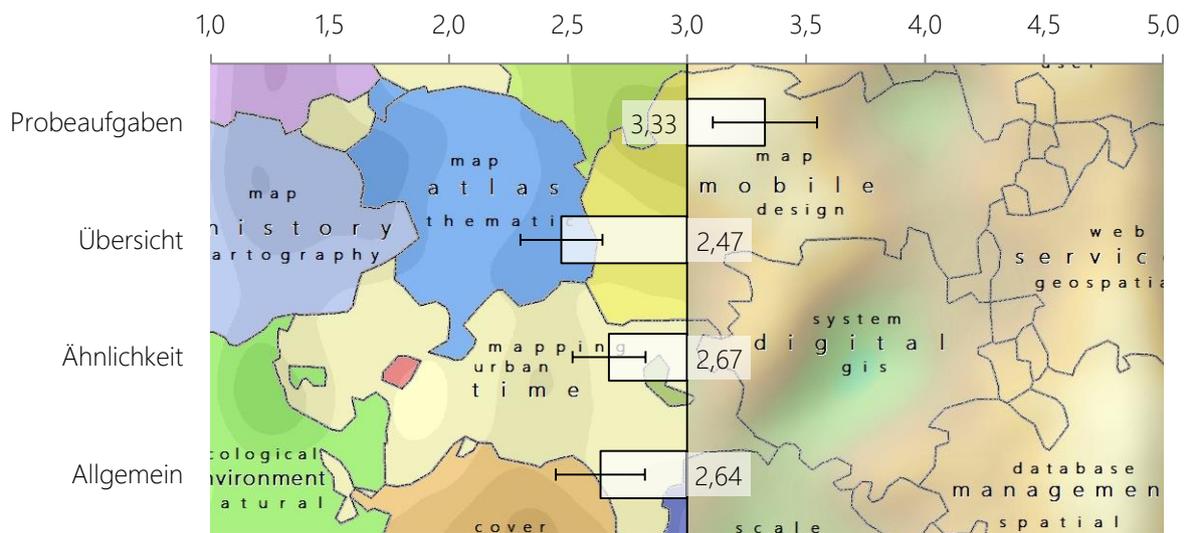


Abbildung 4-11: Präferenz der Testpersonen für einen Visualisierungstyp bei unterschiedlichen Fragetypen [n=55] - Durchschnitt und  $\pm SF$

Wie abgebildet, variiert die Präferenz je nach Fragestellung. Vergleicht man die Bewertung der Präferenz für die drei Fragentypen und allgemein, so unterscheidet sich die Favorisierung signifikant [ $F(3,159)=4.177$ ,  $p<0.05$ ;  $\eta_p^2=0.073$ ].

### 4.3 GEOGRAPHISCHE AUSBILDUNG

Zusätzlich zu den Visualisierungstypen vergleicht man, ob eine geographische Ausbildung die Wahrnehmung einer solchen Informationslandschaft verändert. Wie schon im Kapitel 3.5.4.4 beschrieben, werden dabei nur die Hauptfach-Geographie-Studierende mit den anderen Testpersonen verglichen. Dabei ist dieses Kapitel wieder gleich aufgebaut wie das vorherige: Zuerst werden die Übersichtsaufgaben verglichen und dann die Ähnlichkeitsaufgaben. Die dargestellte Linie in den Diagrammen stellt jeweils den linearen Unterschied zwischen den beiden Stichprobengruppen für die einzelnen Visualisierungstypen dar (violett: Visualisierungstyp D; grün: Visualisierungstyp K).

#### 4.3.1 ÜBERSICHT

Wie man in Abbildung 4-12 erkennt, ist der festgestellte Unterschied in Kapitel 4.2.1 zwischen den beiden Visualisierungstypen bei der geographischen Ausbildung noch ausgeprägter. Die Fragen wurden im Schnitt auch besser von den Hauptfach-Geographie-Studierenden beantwortet. Dennoch hat die geographische Ausbildung keinen signifikanten Einfluss auf das Beantworten der Übersichtsaufgaben [ $F(1,58)=0.082, p>0.05; \eta_p^2=0.001$ ].

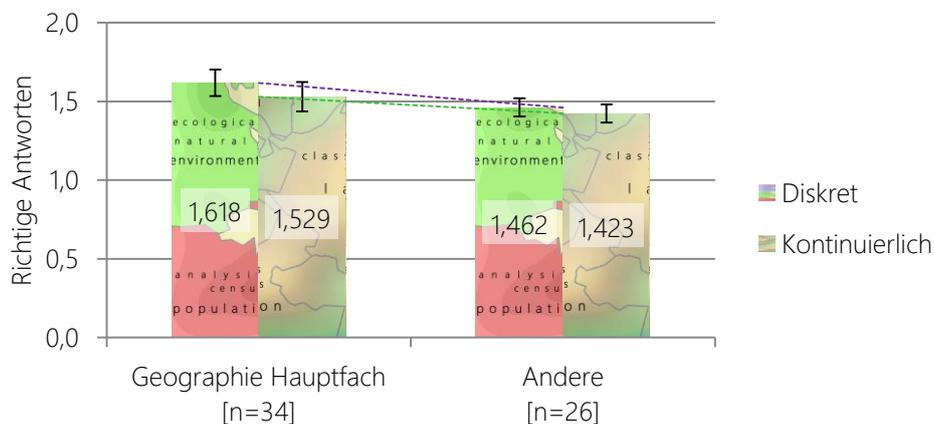


Abbildung 4-12: Vergleich der Ausbildung bei Übersichtsaufgaben - Durchschnitt und ±SF

Zusätzlich zu der Performance kann man das von den Testpersonen eingeschätzte Sicherheitsgefühl vergleichen (Abbildung 4-13). Die Skala reicht wie gehabt von 1 («Sehr unsicher») bis 6 («Sehr sicher»). Der Unterschied zwischen den Visualisierungstypen ist wiederum nicht signifikant [ $F(1,58)=0.693, p>0.05; \eta_p^2=0.012$ ]. Das Geographie-Studium hingegen hat einen bedeutenden Einfluss auf das Sicherheitsgefühl beim Beantworten dieser Informationslandschaften [ $F(1,58)=8.147, p<0.05; \eta_p^2=0.123$ ].

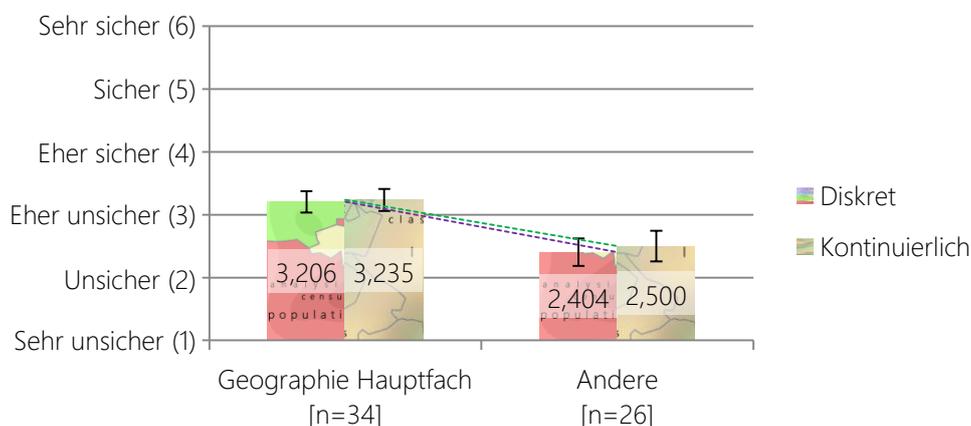


Abbildung 4-13: Vergleich der Ausbildung für die Sicherheit beim Beantworten der Übersichtsaufgaben - Durchschnitt und ±SF

### 4.3.2 ÄHNLICHKEIT

Wie schon im generellen Vergleich analysiert man, ob die Differenzen der Ähnlichkeitswahrnehmung abhängig von der Ausbildung sein können. Dabei erkennt man in Abbildung 4-14 dasselbe Resultat wie in Abbildung 4-7: bei der kontinuierlichen Darstellung schätzte die durchschnittliche Testperson die räumliche Distanz ein bisschen wichtiger ein als bei der diskreten. Derselbe Unterschied gilt zwischen den Geographinnen und Geographen und den Nicht-Geographinnen und Nicht-Geographen.

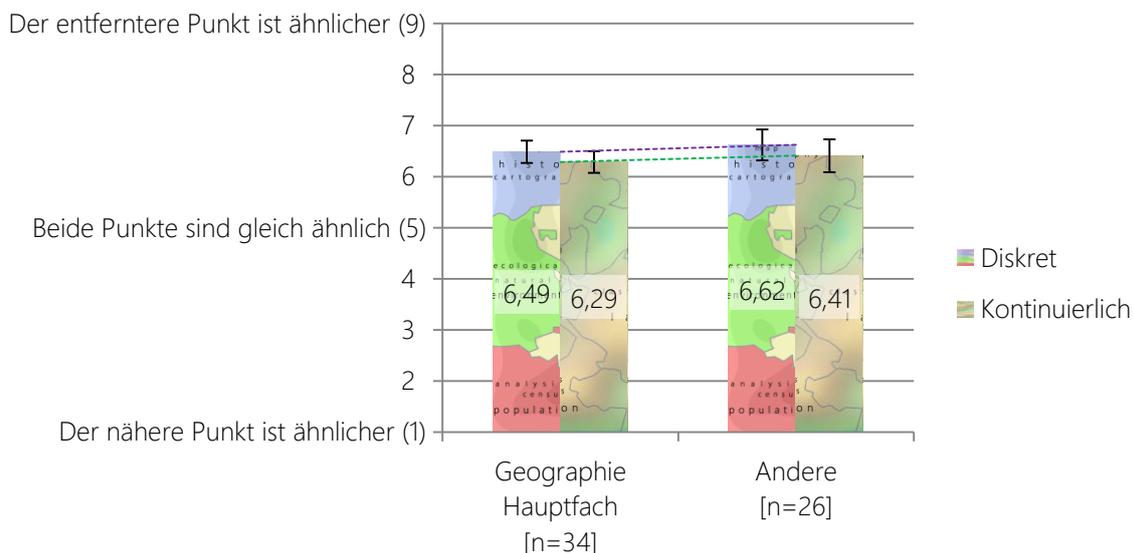


Abbildung 4-14: Vergleich der Ausbildung für die Ähnlichkeitswahrnehmung - Durchschnitt und ±SF

Bei diesem Vergleich nutzten die Geographinnen und Geographen mehr die räumliche Nähe als Metapher für die Ähnlichkeit. Trotzdem ist dieser Unterschied nicht signifikant [F(1,58)=0.134, p>0.05;  $\eta_p^2=0.002$ ]. Und wenn man wieder die Auftrennung in richtig oder falsch bei einem Ähnlichkeitswert von 5 macht, dann erhält man eine Matrix (Tabelle 4-4), welche den Anteil an richtigen Antworten für den jeweiligen Visualisierungstyp mit der Ausbildung der Testperson in Verbindung stellt.

Tabelle 4-4: Matrix der korrekten Antworten der Ähnlichkeitsfragen

	Visualisierungstyp K	Visualisierungstyp D	Total
<b>Geographie Hauptfach</b>	20,1%	17,6%	18,9%
<b>Andere</b>	18,0%	18,6%	18,3%
<b>Total</b>	19,1%	18,1%	18,6%

Dabei kann man schon erkennen, dass die Unterschiede gering sind und man deshalb keine Signifikanz erwartet. Dies wird mit der statistischen Analyse bestätigt [F(1,58)=0.015, p>0.05;  $\eta_p^2=0.000$ ].

Dennoch kann man nun unter den Einflüssen der verschiedenen Faktoren für die Ähnlichkeitswahrnehmung überprüfen, ob die Geographinnen und Geographen auch die räumliche Distanz als einflussreicher eingestuft haben als die Nicht-Geographinnen und Nicht-Geographen. Der allgemeine Schnitt zwischen dem Geographie-Studium und anderen Ausbildungen ist nicht signifikant unterschiedlich [F(1,58)=1.671, p>0.05;  $\eta_p^2=0.028$ ]. Macht man den Vergleich hierarchisch eine Stufe tiefer auf die einzelnen Faktoren gibt es schon auffälligere Differenzen (Abbildung 4-15).

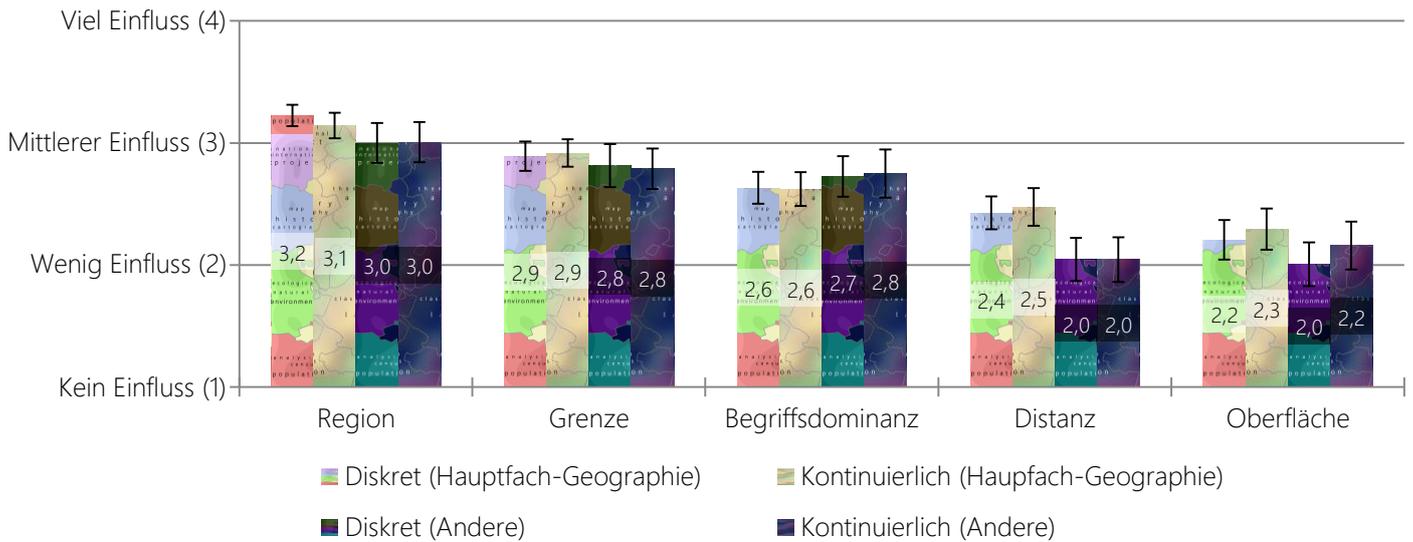


Abbildung 4-15: Vergleich der Ausbildung auf den Einfluss der einzelnen Faktoren- Durchschnitt und ±SF

Betrachtet man nun die statistische Auswertung, gibt es keine signifikanten Wahrnehmungsunterschiede abhängig von einer geographischen Ausbildung. Einzig die Distanz wird von Geographinnen und Geographen deutlich wichtiger eingeschätzt als von Nicht-Geographinnen und Nicht-Geographen. Dieser Unterschied ist bei einem Signifikanzniveau von 0.1 statistisch signifikant.

Tabelle 4-5: Statistische Auswertung der einzelnen Faktoren

Regionszugehörigkeit	F(1,58)=1.100	p>0.05	$\eta_p^2=0.019$
Regionsgrenzen	F(1,58)=0.283	p>0.05	$\eta_p^2=0.005$
Begriffsdominanz	F(1,58)=0.266	p>0.05	$\eta_p^2=0.005$
Distanz	F(1,58)=3.314	0.1>p>0.05	$\eta_p^2=0.054$
Oberflächenbeschaffenheit	F(1,58)=0.471	p>0.05	$\eta_p^2=0.008$

Zuletzt wird auch bei den Ähnlichkeitsaufgaben die Sicherheitseinschätzung untersucht. Dabei resultiert dasselbe Ergebnis wie bei den Übersichtsaufgaben. Die Geographinnen und Geographen beantworten die Fragen im Schnitt sicherer als die Nicht-Geographinnen und Nicht-Geographen. Dieser Unterschied ist signifikant [F(1,58)=6.753, p<0.05;  $\eta_p^2=0.104$ ].

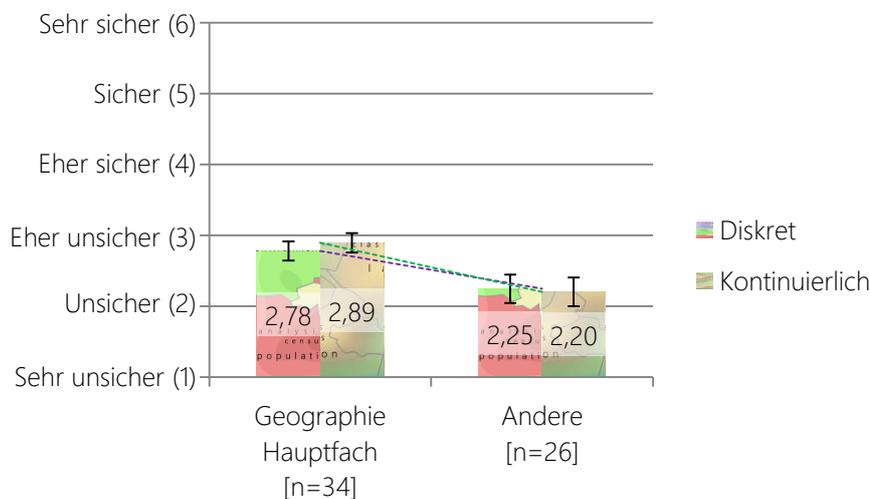


Abbildung 4-16: Vergleich der Ausbildung für die Sicherheit beim Beantworten der Ähnlichkeitsaufgaben - Durchschnitt und ±SF

### 4.3.3 PRÄFERENZ

Vergleicht man nun die Präferenz aufgeteilt nach Studienrichtung, sieht man dabei einige Unterschiede. Die vertikale Achse schneidet die horizontale Achse bei 3 auf einer Skala von 1 («Ziehe diskrete Visualisierung am meisten vor») bis 5 («Ziehe kontinuierliche Visualisierung am meisten vor»). Die ungleiche Favorisierung des Visualisierungstyps je nach Fragestellung ist in Abhängigkeit einer geographischen Ausbildung signifikant [ $F(1,53)=15.065$ ,  $p<0.05$ ;  $\eta_p^2=0.221$ ].

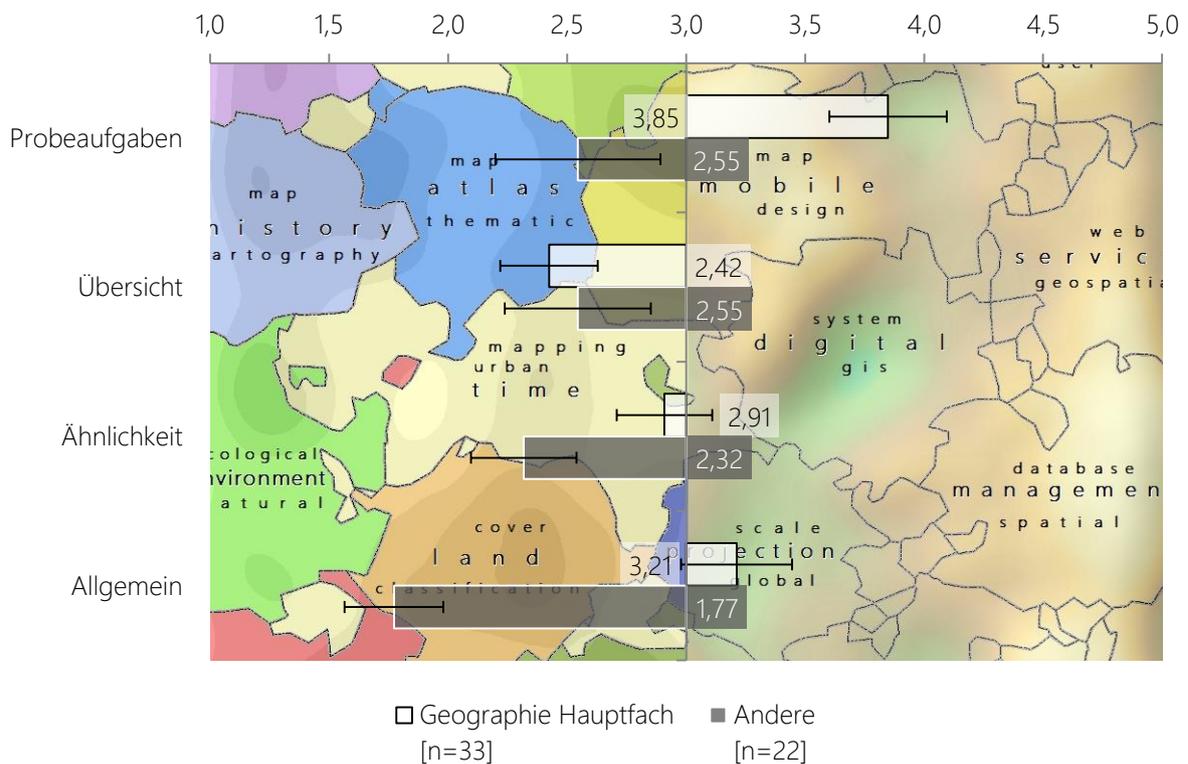


Abbildung 4-17: Präferenz der Testpersonen für einen Visualisierungstyp bei unterschiedlichen Fragetypen (Aufgeteilt nach Ausbildung) - Durchschnitt und ±SF

#### 4.4 ZUSAMMENFASSUNG DER RESULTATE

Überraschenderweise gibt es in der Performance der Teilnehmerinnen und Teilnehmer zwischen den Visualisierungstypen keinen signifikanten Unterschied. Die Fragen zur Übersicht wurden nicht signifikant besser mit der diskreten Visualisierung beantwortet, obwohl sich die Testpersonen bei der diskreten Visualisierung sicherer fühlten. Der einzige signifikante Unterschied zwischen den Visualisierungen ist der Einfluss der Oberflächenbeschaffenheit bei den Ähnlichkeitsaufgaben. Diese wird bei der kontinuierlichen Visualisierung als einflussreicher erachtet. Obwohl die weiteren Faktoren keine signifikante Differenz aufweisen, kann man kleine Unterschiede erkennen. Die Regionszugehörigkeit hat bei der diskreten Visualisierung einen grösseren Einfluss als bei der kontinuierlichen – im Kontrast dazu steht die räumliche Distanz. Beim Beantworten der Fragen mit der kontinuierlichen Visualisierung fühlten sich die Testpersonen auch tendenziell sicherer. Über alle in dieser Studie ausgewerteten Aufgaben hinweg wurde die diskrete Visualisierung präferiert. Einzig bei den Probeaufgaben, in welchen die Testpersonen einzelne Konferenzbeiträge anhand der Begriffsdominanz markieren mussten, wurde die kontinuierliche Darstellung favorisiert.

Neben den beiden Visualisierungstypen möchte diese Arbeit auch den Einfluss einer geographischen Ausbildung untersuchen. Bei den Übersichtsaufgaben wurden die Aufgaben von Geographinnen und Geographen im Schnitt besser gelöst, dennoch besteht keine Signifikanz. Signifikant dagegen ist der Einfluss des Geographie-Studiums bei der Sicherheit im Beantworten der Fragen. Hierbei fühlen sich die Hauptfachstudentinnen und- Studenten tendenziell sicherer. Vergleicht man nun die Ähnlichkeitsaufgaben, fällt auf, dass der durchschnittliche Geograph beziehungsweise die durchschnittliche Geographin wiederum eher den näheren Punkt als ähnlicher bewertet hat. Dieser Zusammenhang ist zwar wiederum nicht signifikant, lässt sich aber bei den eingeschätzten Faktoren näher untersuchen. Tatsächlich schreiben die Geographinnen und Geographen der räumlichen Distanz signifikant eine höhere Bedeutung zu als die Nicht-Geographinnen und Nicht-Geographen. Auch die Regionszugehörigkeit und die Oberflächenbeschaffenheit sind mit einem geographischen Studium ein wenig wichtiger als mit anderen Ausbildungen. Das Sicherheitsgefühl beim Beantworten dieser Fragen ist wiederum bei Geographinnen und Geographen signifikant besser als bei anderen Ausbildungen. Bei der Präferenz gehen die Meinungen zwischen den beiden Stichproben extrem auseinander. Während die Nicht-Geographinnen und Nicht-Geographen konsequent die diskrete Visualisierung bevorzugen, präferieren die Geographinnen und Geographen für die Probeaufgaben sowie allgemein die kontinuierliche Visualisierung.

---

## 5 DISKUSSION

---

In diesem Kapitel werden die Fragestellungen dieser Arbeit im Zusammenhang mit den im vorherigen Kapitel beschriebenen Resultaten und dem aktuellen Forschungsstand analysiert und beantwortet. Beendet wird die Diskussion mit dem Aufzeigen von Verbesserungsmöglichkeiten der Informationslandschaften sowie den Einschränkungen und Problemen der durchgeführten Studie. Da eine der Forschungsfragen eine Differenzierung der Stichprobe anhand der geographischen Ausbildung verlangt, wird die dazugehörige Diskussion der Resultate im jeweiligen Unterkapitel geführt.

### 5.1 LEISTUNG

Wie die Fragestellung auch teilt sich dieses Kapitel in einzelne Sektionen, welche einen Einblick in die Leistung der einzelnen Testpersonen geben sollen. Zur Leistung gehören neben der Korrektheit der Übersichtsaufgaben auch die Bewertung der Ähnlichkeit und das Beurteilen des Einflusses der einzelnen Faktoren auf die Antwort in den Ähnlichkeitsaufgaben. Zusätzlich soll ausgewertet werden, ob sich die Testpersonen bei einem der beiden Visualisierungstypen signifikant sicherer gefühlt haben.

#### 5.1.1 PERFORMANCE

##### 5.1.1.1 Übersichtsaufgaben

Das geographische Konzept der Anordnung soll mit diesem Aufgabentyp für die beiden Visualisierungstypen analysiert werden. Dabei interessiert die Wahrnehmung der Punktgruppierung oder deren Streuung (Fabrikant 2000). In der Studie ist kein signifikanter Unterschied beim korrekten Beantworten der Aufgaben zwischen den Visualisierungstypen festgestellt worden. Die Antworten wurden also unabhängig vom Visualisierungstyp gelöst. Die Übersichtsaufgaben wurden mit 77% zu 74% richtige Antworten besser mit der diskreten Visualisierung bearbeitet. Der Grund für diesen nicht signifikanten aber dennoch erwarteten Unterschied ist eine vereinfachte Zuordnung der Punkte zu den einzelnen Regionen durch den gemeinsamen qualitativen Farbton in der diskretisierten Visualisierung und der durch die Farbe entstandene Verstärkung der Gruppenwahrnehmung. So haben Atkinson et al. (1976) herausgefunden, dass eine Unterteilung der Punkte mit Farbe bei Abschätzungen zu total weniger Fehlern führen. Aber auch die Form und Grösse der zu bearbeitenden Flächen haben einen Einfluss auf die Antworten. Da nach der absoluten Anzahl der Konferenzbeiträge gefragt wurde und deren Punkte in den Regionen verteilt sind, wirkt eine kleine Region schneller überfüllt als eine grosse Region (Lewandowsky et al. 1993). Es wurde auch explizit in den Kommentaren erwähnt, dass die einzelnen farbigen Regionen die Übersicht der weissen Punkte als Konferenzbeiträge in der diskreten Darstellung verbessert. Dies macht durchaus Sinn, da auch bei gewöhnlichen politisch-gegliederten Karten die räumliche Anordnung im Zentrum steht. Für die weitere Diskussion dieses Fragetyps fehlen bereits durchgeführte Experimente und Studien ähnlicher Art, die man mit den Ergebnissen dieser Studie vergleichen könnte. Trotzdem kann man sagen, dass die Erwartungen zwar erfüllt wurden und mit dem Visualisierungstyp D im Schnitt mehr Fragen korrekt beantwortet wurden, aber diese Differenz nicht signifikant ist. Der Visualisierungstyp hat also keinen Einfluss auf das Beantworten der Übersichtsaufgaben.

### 5.1.1.2 Ähnlichkeit

Vergleicht man nun die Ähnlichkeitswahrnehmung der beiden Visualisierungstypen wird wie bei Fabrikant et al. (2006) zuerst aufgrund der Region als geographische Metapher entschieden. So wurde, unabhängig des Visualisierungstyps oder der Fragestellung, im Schnitt immer der distanziertere Punkt als ähnlicher erachtet, wenn er in derselben Region liegt. Die räumliche Distanz zwischen den Konferenzbeiträgen und die als Topographie dargestellte Begriffsdominanz haben wenig Einfluss auf die Ähnlichkeitsbewertung. Trotzdem wurde beim Visualisierungstyp D die Antwort noch mehr von der farblichen Verwandtschaft zu einem Cluster beeinflusst und somit dem Punkt in derselben Region immer eine höhere Ähnlichkeit zugeschrieben als beim Visualisierungstyp K. Denn die untergeordnete Struktur der Punkte in einer Region verzerrt systematisch die Wahrnehmung der räumlichen Beziehungen (Stevens & Coupe 1978) und durch die farbliche Unterscheidung wird diese im Visualisierungstyp D betont. In der einzigen Aufgabe, in welcher die Region als geographische Metapher irrelevant ist, da alle drei bearbeiteten Punkte in derselben Region liegen, wird, bevor man auf die Distanz als Entscheidungsgrundlage zurückgreift, die Begriffsdominanz priorisiert. Dies ist ausserdem die einzige Aufgabe, in welcher ein statistisch signifikanter Unterschied zwischen dem Visualisierungstyp in der Ähnlichkeitswahrnehmung besteht. Dabei spielt beim Visualisierungstyp D die Begriffsdominanz eine wichtigere Rolle als beim Visualisierungstyp K, bei welchem öfter der nähere Punkt als ähnlicher erachtet wird. Vergleicht man nun die von den Testpersonen selbsteingeschätzten Faktoren, welche die Entscheidung jeweils beeinflusst haben, dann bestätigt sich das Resultat in der Ähnlichkeitswahrnehmung. Der Einfluss der Region wird jeweils auch als am einflussreichsten bewertet – die Region ist bei der diskreten Visualisierung sogar noch ein wenig einflussreicher als bei der kontinuierlichen. Dies kann man mit dem Hervorheben der Regionen durch verschiedene Farbtöne begründen. Liegen aber beide Punkte in derselben Region, wurde auch wiederum bei beiden Darstellungstypen die Begriffsdominanz als wichtiger erachtet als die räumliche Distanz. Einzig die Oberflächenbeschaffenheit wurde bei der kontinuierlichen Darstellung als signifikant einflussreicher eingestuft als bei der diskreten Visualisierung. Dies kann man mit der Vertrautheit der Darstellung als Reliefkarte begründen, in welcher auch der Oberflächenverlauf im Mittelpunkt steht. Der Visualisierungstyp beeinflusst zwar nicht direkt das Beantworten der Ähnlichkeitsaufgaben, dennoch wird die Oberflächenbeschaffenheit als Entscheidungsgrundlage beim Visualisierungstyp K als signifikant einflussreicher eingeschätzt. Die Aufmerksamkeit liegt bei diesem Visualisierungstyp auf den Landschaftsformen wie Hügel oder Täler und nicht auf den Regionen (Montello et al. 1994).

Weder bei den Übersichts- noch bei den Ähnlichkeitsaufgaben gibt es einen statistisch signifikanten Unterschied in der Performance zwischen den Visualisierungstypen. Trotzdem entspricht das Resultat den Erwartungen: Bei den Übersichtsaufgaben wurden mehr Fragen mit dem Visualisierungstyp D, bei den Ähnlichkeitsaufgaben mit dem Visualisierungstyp K korrekt beantwortet. Zusätzlich weist die räumliche Distanz im Visualisierungstyp K bei der Ähnlichkeitswahrnehmung einen grösseren Einfluss auf als beim Visualisierungstyp D. Dies konnte erwartet werden, da räumliche Landschaften in Reliefkarten ähnlich dargestellt werden und in diesen das Schätzen der Distanz einen bedeutenden Nutzen darstellt. Dennoch ist nur der selbst eingeschätzte Einfluss der Oberflächenbeschaffenheit für die Ähnlichkeitswahrnehmung beim Visualisierungstyp K signifikant grösser. Dies kann man auch wieder mit der Analogie zur Reliefkarte begründen.

### 5.1.2 RÜCKSCHLÜSSE ZUR ERSTEN FORSCHUNGSFRAGE

Zusammenfassend wird die erste Forschungsfrage so beantwortet: Der Visualisierungstyp hat keinen signifikanten Einfluss auf die Leistung der Testpersonen. Die Hypothese  $H_{(1.1)}$  «*Der Visualisierungstyp beeinflusst die Leistung beim Beantworten der gestellten Aufgaben*» muss also abgelehnt werden. Dennoch hat jeder der beiden Visualisierungstypen abhängig der Fragestellung seine Vorteile. Während die räumliche Anordnung, welche in den Übersichtsaufgaben eine Rolle spielt, beim Visualisierungstyp D betont wird, kommt beim Visualisierungstyp K sowohl die Informationsoberfläche als Topographie als auch die Distanz, die in den Ähnlichkeitsaufgaben bedeutend ist, mehr zur Geltung.

### 5.1.3 SICHERHEIT

Sowohl bei den Übersichtsaufgaben als auch bei den Ähnlichkeitsaufgaben schätzten die Testpersonen im Schnitt ihre Sicherheit beim Beantworten der Frage bei der kontinuierlichen Darstellung höher ein. Bei beiden Aufgabentypen fühlten sich die Testpersonen bei der kontinuierlichen im Schnitt sicherer. Allerdings ist dieser Unterschied nicht signifikant. Zusätzlich muss erwähnt werden, dass die Probandinnen und Probanden bei den Ähnlichkeitsaufgaben durchschnittlich unsicherer waren als bei den Übersichtsaufgaben.

Die Erwartungen wurden also nicht ganz erfüllt. Man ging davon aus, dass sich die Testpersonen bei den Übersichtsaufgaben sicherer bei dem Visualisierungstyp D fühlten. Begründet wird diese Erwartung mit dem hierarchischem Schema (Stevens & Coupe 1978), welches bei diesem Visualisierungstyp mehr zur Geltung kommt. Bei den Ähnlichkeitsaufgaben erwartete man, dass sich die Probandinnen und Probanden mit dem Visualisierungstyp K besser zurechtfinden, was zwar dem Resultat entspricht, aber statistisch auch nicht signifikant ist.

### 5.1.4 RÜCKSCHLÜSSE ZUR ZWEITEN FORSCHUNGSFRAGE

Auch die Beantwortung der zweiten Forschungsfrage wird mit keinem signifikanten Unterschied in der Sicherheitsbewertung begründet. Die Hypothese  $H_{(1.2)}$  «*Der Visualisierungstyp beeinflusst das Sicherheitsgefühl beim Beantworten der gestellten Aufgaben*» muss ebenfalls abgelehnt werden. Es zeigt sich, dass unabhängig vom Aufgabentyp die Probandinnen und Probanden eher unsicher bei der Beantwortung der Fragen waren. Ganz wenig besser hat der Visualisierungstyp K abgeschlossen, welcher im Durchschnitt eine höhere Sicherheit als der Visualisierungstyp D übermittelte.

## 5.2 PRÄFERENZ

Die Präferenz unterscheidet sich je nach Aufgabentyp signifikant. Während einzig für die Probeaufgaben die kontinuierliche Informationslandschaft präferiert wurde, wird die diskrete Darstellung bei den anderen Aufgabentypen und auch generell lieber bearbeitet. Vor allem bei den Übersichtsaufgaben wird die diskrete Informationslandschaft bevorzugt, obwohl sich die Testpersonen bei dieser unsicherer fühlten. Wie Rosch (1973) schon erwähnt hat, ist es angebracht, wenn Leute mit dem Kontinuum eines Phänomens kognitive Mühe haben, dieses diskret darzustellen. Dies zeigt sich in der resultierenden Präferenz wieder. Die Bevorzugung des Visualisierungstyp K bei den Probeaufgaben kann man mit der durchgehenden Darstellung der Begriffsdominanz im Visualisierungstyp K und mit der Unabhängigkeit der Aufgabenstellung von der Regionszuweisung begründen. So kann man dabei die einzelnen Punkte nur anhand des Farbtönen überregional vergleichen und muss nicht wie beim Visualisierungstyp D die Farbtöne und den Farbwert gegenüberstellen.

### 5.2.1 RÜCKSCHLÜSSE ZUR DRITTEN FORSCHUNGSFRAGE

Die Präferenz für eine der beiden Visualisierungstypen, welche in der dritten Forschungsfrage abhängig von der Aufgabe gesucht wird, ist signifikant unterschiedlich. Die Hypothese  $H_{(2)}$  «*Die Testpersonen präferieren einen Visualisierungstyp abhängig von der gestellten Aufgabe*» wird also angenommen. Ein bisschen überraschend ist die Favorisierung des Visualisierungstyp D bei praktisch allen Aufgaben schon. Denn obwohl das diskret dargestellte Phänomen der Begriffsdominanz das Verständnis zwar erleichtert, ist diese Darstellungsart im Gegensatz zum Visualisierungstyp K weniger geläufig.

## 5.3 HINTERGRUNDWISSEN

### 5.3.1 ÜBERSICHTSAUFGABEN

Nun kann man die durchgeführten Vergleiche auf die beiden Stichproben aufteilen, welche den Einfluss des geographischen Hauptfach-Studiums untersucht. Beginnt man mit der Diskussion der Resultate erneut bei den Übersichtsaufgaben, dann erkennt man bei diesem Aufgabentyp, dass die Geographinnen und Geographen die Fragen im Schnitt zwar besser beantwortet haben, dieser Unterschied aber nicht signifikant ist. Beide Stichproben waren mit dem Visualisierungstyp D erfolgreicher. Ein statistisch signifikanter Unterschied besteht zwischen dem Sicherheitsgefühl der Stichprobe mit im Hauptfach Geographie ausgebildeten Personen und der anderen Gruppe. Die Geographinnen und Geographen schätzten ihr Sicherheitsgefühl bei der Beantwortung der Übersichtsfragen statistisch signifikant höher ein. Dies kann man mit der Vertrautheit mit solchen Darstellungen begründen (Moore & Bricker 2015).

### 5.3.2 ÄHNLICHKEIT

Vergleicht man die Ähnlichkeitsaufgaben anhand der Kontrollvariablen der geographischen Ausbildung, dann fallen folgende Resultate auf. Obwohl die geographische Ausbildung bei beiden Visualisierungstypen zu einer höheren Gewichtung des näheren Punktes als bei einem sonstigen Hintergrund führt, weist dieser Unterschied keine statistische Signifikanz auf. Statistisch signifikant ist allerdings der eingeschätzte Einfluss der Distanz zwischen den Stichprobengruppen. So bewerten die Geographinnen und Geographen die Distanz als bedeutend einflussreicher auf ihre Antwort als die andere Stichprobengruppe. Begründen kann man das mit dem vertieftem Auseinandersetzen mit dem ersten Gesetz von Tobler (1970) während dem geographischen Studium. Weitere kleinere Unterschiede bei der Oberfläche oder der Region, welche auch Personen mit einem geographischen Hintergrund als einflussreicher betrachten, weisen keine statistische Signifikanz auf. Die Geographinnen und Geographen schätzten ihr Sicherheitsgefühl bei der Beantwortung der Fragen statistisch signifikant höher ein. Dies begründet man wiederum mit der unterschiedlich häufigen Konfrontation mit solchen Darstellungen (Moore & Bricker 2015).

### 5.3.3 PRÄFERENZ

Vergleicht man zusätzlich die Präferenzen verbunden mit einer geographischen Ausbildung ergibt sich ein völlig anderes Bild. Die Nicht-Geographinnen und Nicht-Geographen favorisieren, unabhängig von der Fragestellung, die diskrete Visualisierung und die Geographinnen und Geographen verwenden für die Probeaufgaben und auch allgemein lieber die kontinuierliche Darstellung. Bei den Ähnlichkeitsaufgaben bearbeiten die Testpersonen unabhängig von der geographischen Vertiefung beide Visualisierungstypen etwa gleich gerne.

### 5.3.4 RÜCKSCHLÜSSE ZUR VIERTEN FORSCHUNGSFRAGE

Der Einfluss eines Geographie-Studiums wird in der vierten Forschungsfrage untersucht. Der Unterschied zwischen Testpersonen mit einem geographischen Hintergrund und den Übrigen ist statistisch signifikant. Die Hypothese  $H_{(3)}$  «Eine geographische Ausbildung beeinflusst die Leistung beim Beantworten der gestellten Aufgaben» wird also angenommen. Zusammenfassend kann man also behaupten, dass die geographische Ausbildung die Deutung der räumlichen Distanz in der Leistung zwar beeinflusst, allerdings nicht statistisch signifikant. Bei der Einschätzung der verschiedenen Faktoren ist die Distanz allerdings für die Geographinnen und Geographen signifikant einflussreicher. Ausserdem führt eine geographische Ausbildung auch zu einem erhöhten Sicherheitsgefühl beim Beantworten der Fragen. Die Studie bestätigt die Resultate für die Distanz-Ähnlichkeitsmetapher mit Regionen in der Publikation von Fabrikant, Montello & Mark (2006).

## 5.4 VERBESSERUNGSMÖGLICHKEITEN FÜR INFORMATIONSLANDSCHAFTEN

Ganz ausgereift sind die in dieser Arbeit erstellten Informationslandschaften nicht. Es gibt einige Verbesserungsmöglichkeiten, die allerdings im Rahmen dieser Arbeit nicht mehr angewandt werden konnten. Einzelne Kritikpunkte haben auch schon Studienteilnehmerinnen und –Teilnehmer in den Kommentaren geäußert. Aufgefallen ist, dass bei der kontinuierlichen Visualisierung die Begriffsdominanz je höher desto heller dargestellt wird. Bei der diskreten Darstellung nimmt die Helligkeit mit der Begriffsdominanz ab. Begründen kann man allerdings beide Darstellungsarten. Die kontinuierliche Visualisierung basiert auf den Höhenstufen vom topographischen Gelände nach Imhof (1965). Analog dieser Norm erscheinen die Hügel und Berge in einem helleren Farbton. Die klassierte Darstellung der Begriffsdominanz basiert auf einer sequentiellen Farbwahl, in welcher die Begriffsdominanz mit dem Farbwert zunimmt. Es bezieht sich also auf die qualitative Unterscheidung der Regionen mit dem Farbton und die quantitative Differenzierung der Begriffsdominanz mit dem Farbwert (Bertin 1967).

Eine weitere Änderung, welche eher die Datengrundlage widerspiegelt, wäre wie bei Biberstine et al. (2010) oder Skupin & de Jongh (2005) die Beschriftung der einzelnen Begriffsdominanzhöhen. Dabei werden nicht die Regionen, sondern immer jeweils die häufigsten Begriffe pro Neuron in der Informationslandschaft beschriftet und von der Region ist nur die jeweilige Grenze zu sehen. Im Unterschied zu dieser Methode wird in dieser Arbeit die Region analog zu Ländern in einer thematischen Karte beschrieben. Eine zusätzliche Erweiterung der verbalen Datenrepräsentation bietet das sogenannte *Stemming* der Wörter. Dieses algorithmische Verfahren führt verschiedene morphologische Varianten eines Wortes auf ihren Wortstamm zurück (Porter 1980). So könnte die Häufigkeit für den jeweiligen Wortstamm kalkuliert werden und so die Vielfalt der verschiedenen Wortformen auf einen Begriff reduzieren.

Die Beschriftung stellte sich als grosse Herausforderung dar. Aufgrund der fokussierten Datengrundlage ist eine solche wissenschaftliche Differenzierung der in der Erkennung von *Blondel Communités* extrahierten Clusteranzahl von 18 kompliziert. Daher wurde eine Umfrage mit den TF/IDF-Werten der einzelnen Cluster in einem kleineren Rahmen durchgeführt, welche aber die Befragten schon ziemlich strapazierte. Diese Unterscheidung ist mit einer breiteren Datengrundlage in den Daten wie bei Skupin (2010), Biberstine et al. (2012) oder Skupin et al. (2013) einfacher.

Ein weiterer Kritikpunkt in den Ausgangsdaten ist das Verwenden von verschiedenen Informationsmengen. So wurden in dieser Arbeit auch Konferenzbeiträge verwendet, welche nur einen englischen Titel aufweisen. Bei einer solchen Ausgangslage wird das Topic Modeling – also die Zuweisung und Aufteilung der Themen zu einem solchen Beitrag – ein bisschen verzerrt. Damit sich die Beiträge verhältnismässiger und thematisch sinnvoller in der Informationslandschaft verteilen müsste man die Informationsmenge für alle Beiträge homogenisieren. Denn gemäss Steyvers & Griffiths (2007) ist einzig die Häufigkeit der einzelnen thematischen Wörter pro Artikel die entscheidende Grundlage für die aus dem Topic Modeling resultierende Artikel-Themen-Matrix. Dies wurde in dieser Informationslandschaft vernachlässigt, da ansonsten die erste verfügbare Konferenz (1993) in Köln mit nur 66 Abstracts gegenüber 138 Beiträgen (Titel und/oder Abstracts) beinahe ausgeschlossen wird.

## 5.5 LIMITIERUNGEN UND PROBLEME DER STUDIE

Eines der wahrscheinlich einflussreichsten Probleme ist die physische Abwesenheit einer Ansprechperson. Trotz der ausführlichen Anleitung und der Angabe der Kontaktinformationen, welche nur selten für Unklarheiten genutzt wurden, scheinen vor allem Personen, welche so eine Darstellung zum ersten Mal gesehen haben, Mühe mit dem Verständnis zu haben. Dies hat einen durchgehenden Einfluss auf die Antworten der Studie.

Zusätzlich könnte man dann die Zeit für das Beantworten der Fragen als abhängige Variable verwenden. Dies würde vor allem bei den Übersichtsaufgaben einen verbesserten Vergleich der Visualisierungstypen ermöglichen. Für eine vertiefte Auswertung der Übersichtsaufgaben wäre es ausserdem sinnvoll, wenn genau dieselbe Fragestellung für beide Visualisierungstypen bei diesem Aufgabentyp bearbeitet werden muss. Dies wurde nicht gemacht, da man einen Lerneffekt befürchtet hat und die Wahrscheinlichkeit, dass bei der zweiten Darstellung zum selben Typ ohne Überlegen eine Antwort gewählt wird, gross ist.

Ausserdem war es nicht möglich, eine Aufgabe im Posttest auf einem mobilen Gerät wie einem Smartphone oder Tablet durchzuführen, in welchen verschiedene Strecken anhand der Entfernung sortiert werden mussten. Dies ist der Grund, weshalb bei der Präferenz die Anzahl der Testpersonen um fünf kleiner ist. Neben der Problematik dieser Fragestellung im Posttest hätte man neben der Ähnlichkeit auch die Einflüsse der Faktoren auf die Antwort abfragen müssen. So könnte man den Vergleich zwischen einer Informationslandschaft und einer reellen räumlichen Landschaft machen. Dieser Vergleich ist in dieser Studie leider nicht möglich, da in dem Ausschnitt der europäischen Landschaft die Regionen als hierarchisches Schema fehlen.



## 6 ZUSAMMENFASSUNG & AUSBLICK

---

Wie man aus dem Forschungsstand zu erkennen vermag bildet die Informationslandschaft eine alternative Visualisierungsmethode grosser relationaler Daten. Verschiedene geographische Metaphern wie die Region mit dem dahinterliegenden hierarchischen Schema können problemlos auf die Visualisierung nicht-räumlicher Daten in einem Informationsraum übertragen werden. Diese Metapher wird in der Ähnlichkeitswahrnehmung durchgehend priorisiert. Den weiteren Metaphern wie der Distanz-Ähnlichkeits-Metapher wird weniger Bedeutung zugesprochen. Untersucht man die Ähnlichkeit bei Landschaften ohne diskrete Grenzen, wird vor der räumlichen Distanz die Topographie der Informationsoberfläche beachtet. Diese Priorisierung der geographischen Metaphern ist für beide Visualisierungstypen dieselbe. Dennoch kann man die Darstellungsmethode einer thematischen oder topographischen Karte in einer Informationslandschaft anwenden. Die Differenzen sind zwar nicht signifikant, dennoch erkennt man gewisse Trends. So werden Aufgaben, in welcher die räumliche Anordnung einzelner Elemente wesentlich ist, mit dem Visualisierungstyp D, der einer politisch-gegliederten Karte angelehnt ist, besser gelöst. Bei Aufgaben, in welchen die relativen oder absoluten Positionen einzelner diskreter, überregional verteilter Objekte in der Informationslandschaft relevant sind, dient der Visualisierungstyp K als eine grössere Hilfe.

Die Auswertungen in dieser Arbeit deuten darauf hin, dass der Visualisierungstyp einen kleineren Einfluss als eine geographische Ausbildung auf das Verständnis dieser Informationslandschaften hat. Dabei spielt das erste Gesetz der Geographie von Tobler (1970), welches im Studium eingeprägt wird, eine bedeutende Rolle. So sind solche Informationslandschaften für viele Nicht-Geographinnen und Nicht-Geographen ein Rätsel, welches in dieser Studie hauptsächlich mit der geographischen Metapher der Region gelöst wurde. Wird die Region wie bei der Ähnlichkeitsaufgabe 5 irrelevant gibt es einen signifikanten Unterschied. Deshalb bieten sich weitere Forschungsarbeiten und Studien zu Informationslandschaften ohne Regionen an.



---

## 7 LITERATUR & QUELLEN

---

- ATKINSON, J., FRANCIS, M. R., & CAMPBELL, F. W. (1976). The dependence of the visual numerosity limit on orientation, colour, and grouping in the stimulus. *Perception*, 5(3), S. 335–342.
- BATAGELJ, V., & MRVAR, A. (2009). Pajek64 - Program for Analysis and Visualization of Large Networks. Abgerufen von <http://pajek.imfm.si/> (Letzter Zugriff: Juni 2015)
- BERTIN, J. (1967). *Sémiologie Graphique. Les diagrammes, les réseaux, les cartes*. Paris: Gauthier-Villars.
- BIBERSTINE, J., DUHON, R. J., BÖRNER, K., HARDY, E., & SKUPIN, A. (2010). *A Semantic Landscape of the Last. fm Music Folksonomy: Using a Self-organizing Map*.
- BIBERSTINE, J., DUHON, R. J., BÖRNER, K., HARDY, E., & SKUPIN, A. (2012). A Semantic Map of the last. fm Music Folksonomy. *cns.slis.indiana.edu*.
- BIRD, S., KLEIN, E., & LOPER, E. (2009). *Natural Language Processing with Python*. O'Reilly Media.
- BLEI, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), S. 77–84.
- BLONDEL, V. D., GUILLAUME, J.-L., LAMBIOTTE, R., & LEFEBVRE, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, S. P10008.
- BÖRNER, K., CHEN, C., & BOYACK, K. W. (2003). Visualizing knowledge domains. *Annual Review of Information Science and Technology*, 37(1), S. 179–255.
- BOYACK, K. W., KLAVANS, R., & BÖRNER, K. (2005). Mapping the backbone of science. *Scientometrics*, 64(3), S. 351–374.
- BOYACK, K. W., NEWMAN, D., DUHON, R. J., KLAVANS, R., PATEK, M., BIBERSTINE, J., ET AL. (2011). Clustering more than two million biomedical publications: comparing the accuracies of nine text-based similarity approaches. *PloS one*, 6(3), S. e18029.
- BREWER, C. A. (1994). Color Use Guidelines for Mapping and Visualization, in: A.M. MacEachren and D.R.F. Taylor (Hrsg.), *Visualization in Modern Cartography*, S. 123–147. Tarrytown, NY: Elsevier Science.
- BRUGGMANN, A. (2012). Netzwerkvisualisierung der Ostschweiz. Masterarbeit. Geographisches Institut der Universität Zürich.
- BRUGGMANN, A., SALVINI, M. M., & FABRIKANT, S. I. (2013a). Cartograms of self-organizing maps to explore user-generated content, in: *Proceedings of the 26th International Cartographic Conference*. Dresden, Germany.
- BRUGGMANN, A., SALVINI, M. M., & FABRIKANT, S. I. (2013b). Agglomerationen mit nutzergenerierten Inhalten neu definiert Visualisierung der Nordostschweiz mithilfe von Wikipedia. *disP - The Planning Review*, 49(4), S. 37–45.
-

- BUCKLEY, A. R., GAHEGAN, M., & CLARKE, K. (2000). Emerging Themes in GIScience Research: Geographic Visualization. *University Consortium for Geographic Information Science (UCGIS), White Papers*.
- CARD, S. K., MACKINLAY, J., & SHNEIDERMAN, B. (1999). *Readings in Information Visualization: Using Vision to Think*. San Francisco, CA: Morgan Kaufmann.
- CHALMERS, M. (1993). *Using a Landscape Metaphor to Represent a Corpus of Documents* (S. C. Hirtle & A. U. Frank, Hrsg.). Springer-Verlag Berlin Heidelberg.
- CHEN, H., HSU, P., ORWIG, R., HOOPEES, L., & NUNAMAKER, J. F. (1994). Automatic concept classification of text from electronic meetings. *Communications of the ACM*, 37(10), S. 56–73.
- COUCLELIS, H. (1992). People manipulate objects (but cultivate fields): Beyond the Raster-Vector Debate in GIS. *Theories and Methods of Spatiotemporal Reasoning in Geographic Space*, 639(716), S. 65–77.
- COUCLELIS, H. (1998). Worlds of Information: The Geographic Metaphor in the Visualization of Complex Information. *Cartography and Geographic Information Science*, 25(4), S. 209–220.
- DODGE, M., & KITCHIN, R. (2001). *Mapping Cyberspace*. Routledge Chapman & Hall.
- ESRI. (2014). ArcGIS 10 - Mapping and Analysis of Data. (Letzter Zugriff: September 2015)
- FABRIKANT, S. I. (2000). Spatialized Browsing in Large Data Archives. *Transactions in GIS*, 4(1), S. 65–78.
- FABRIKANT, S. I. (2001). Visualizing region and scale in information spaces, in: *Proceedings of the 20th International Cartographic Conference*, S. 2522–2529. Beijing, China.
- FABRIKANT, S. I., & BUTTENFIELD, B. P. (2001). Formalizing Semantic Spaces For Information Access. *Annals of the Association of American Geographers*, 91(2), S. 263–280.
- FABRIKANT, S. I., MONTELLO, D. R., & MARK, D. M. (2006). The distance-similarity metaphor in region-display spatializations. *Computer Graphics and Applications, IEEE*, 26(4), S. 34–44.
- FABRIKANT, S. I., MONTELLO, D. R., & MARK, D. M. (2010). The natural landscape metaphor in information visualization: The role of commonsense geomorphology. *Journal of the American Society for Information Science and Technology*, 61(2), S. 253–270.
- FABRIKANT, S. I., MONTELLO, D. R., RUOCCO, M., & MIDDLETON, R. S. (2004). The Distance–Similarity Metaphor in Network-Display Spatializations. *Cartography and Geographic Information Science*, 31(4), S. 237–252.
- FABRIKANT, S. I., RUOCCO, M., MIDDLETON, R., MONTELLO, D. R., & JÖRGENSEN, C. (2002). The first law of cognitive geography: distance and similarity in semantic space, in: *Proceedings of GIScience 2002*, S. 31–33.
- FABRIKANT, S. I., & SALVINI, M. M. (2011). Charting the ICA World of Cartography 1999-2009, in: *Proceedings of the 25th International Cartographic Conference*, S. 1–13. Paris, France.

- FABRIKANT, S. I., & SKUPIN, A. (2005). Cognitively Plausible Information Visualization, in: J. Dykes, A. M. MacEachren, & M.-J. Kraak (Hrsg.), *Exploring Geovisualization*, S. 667–690. Amsterdam: Elsevier.
- FEINBERG, J. (2014). Wordle. Abgerufen von <http://www.wordle.net/> (Letzter Zugriff: September 2015)
- FIELD, A. (2013). *Discovering statistics using IBM SPSS statistics*. SAGE Publications Ltd.
- FRIEDMAN, A., & BROWN, N. R. (2000). Reasoning about geography. *Journal of Experimental Psychology: General*, 129(2), S. 193–219.
- GIRAUD, H. (2013). Strasbourg, in: *Proceedings of the 26th International Cartographic Conference*. Dresden, Germany.
- HESPANHA, S. R., & HESPANHA, J. (2011). Text Visualization Toolbox — a MATLAB toolbox to visualize large corpus of documents. (Letzter Zugriff: Juni 2015)
- HUFFMANN, N. H. (1993). Hyperchina: Adventures in hypermapping, in: *Proceedings of the 16th International Cartographic Conference*, S. 26–45. Cologne, Germany.
- IBM. (2012). IBM SPSS - Statistikprogramm. Abgerufen von <http://www-01.ibm.com/software/ch/de/analytics/spss/> (Letzter Zugriff: September 2015)
- IKG ETH ZÜRICH. (2015). SCHWEIZER WELTATLAS interaktiv. Abgerufen von <http://schweizerweltatlas.ch/> (Letzter Zugriff: September 2015)
- IMHOF, E. (1965). *Kartographische Geländedarstellung*. Berlin: De Gruyter.
- ISHIKAWA, T., & KASTENS, K. A. (2005). Why some students have trouble with maps and other spatial representations. *Journal of Geoscience Education*, 53(2), S. 184–197.
- JESSUP, E. R., DRMAC, Z., & BERRY, M. W. (1999). Matrices, vector spaces, and information retrieval. *SIAM review*, 41(2), S. 335–362.
- DE JONGH, C. (2003). Mapping a Cartographic Conference: the Experimental Spatialisation of Non-Spatial Information, in: *Proceedings of the 21st International Cartographic Conference*, S. 2129–2142. Durban, South Africa.
- KEIM, D., PANSE, C., & SIPS, M. (2005). Information visualization: Scope, techniques and opportunities for geovisualization, in: J. Dykes, A. M. MacEachren, & M.-J. Kraak (Hrsg.), *Exploring Geovisualization*, S. 21–52. Amsterdam: Elsevier.
- KLAVANS, R., & BOYACK, K. W. (2006). Identifying a Better Measure of Relatedness for Mapping Science. *Journal of the American Society For Information Science And Technology*, 57(2), S. 251–263.
- KOHONEN, T. (2001). *Self-Organizing Maps*. Springer-Verlag Berlin Heidelberg.
- KOHONEN, T., KASKI, S., & LAGUS, K. (2000). Self organization of a massive document collection. *IEEE Transactions on Neural Networks*, 11(3), S. 574–585.

- KRAAK, M. J. (2003). Geovisualization illustrated. *ISPRS Journal of Photogrammetry and Remote Sensing*, 57(5-6), S. 390–399.
- LACAYO-EMERY, M. (2011). An Integrated Toolset for Exploration of Spatio- Temporal Data Using Self-Organizing Maps and Gis. San Diego, CA: Montezuma Publishing.
- LANDAUER, T. K., FOLTZ, P. W., & LAHAM, D. (1998). An Introduction to Latent Semantic Analysis. *Discourse Processes*, 25, S. 259–284.
- LEWANDOWSKY, S., HERRMANN, D. J., BEHRENS, J. T., LI, S. C., PICKLE, L., & JOBE, J. B. (1993). Perception of clusters in statistical maps. *Applied Cognitive Psychology*, 7(6), S. 533–551.
- MACEachREN, A. M., & KRAAK, M.-J. (2001). Research Challenges in Geovisualization. *Cartography and Geographic Information Science*, 28(1), S. 3–12.
- MACKINLAY, J. (1986). Automating the design of graphical presentations of relational information. *ACM Transactions on Graphics (TOG)*, 5(2), S. 110–141.
- MARTIN, D. W. (2008). *Doing Psychology Experiments*. Cengage Learning.
- MATHWORKS. (2010). MATLAB - Software zur Lösung mathematischer Probleme und zur graphischen Darstellung. Abgerufen von <http://www.mathworks.ch> (Letzter Zugriff: Juni 2015)
- MICROSOFT. (2010). Microsoft Excel. (Letzter Zugriff: September 2015)
- MOELLERING, H., & AHLQVIST, O. (2013). Columbus, in: *Proceedings of the 26th International Cartographic Conference*. Dresden, Germany.
- MONTELLO, D. R., FABRIKANT, S. I., RUOCCO, M., & MIDDLETON, R. S. (2003). Testing the First Law of Cognitive Geography on Point-Display Spatializations, in: *Spatial Information Theory. Foundations of Geographic Information Science*, S. 316–331. Springer-Verlag Berlin Heidelberg.
- MONTELLO, D. R., SULLIVAN, C. N., & PICK., H. L. (1994). Recall Memory for Topographic Maps and Natural Terrain: Effects of Experience and Task Performance. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 31(3), S. 18–36.
- MOORE, A. B., & BRICKER, M. (2015). 'Mountains of work': spatialization of work projects in a virtual geographic environment. *Annals of GIS*, S. 1–11.
- MYSQL. (2011). Open-Source-Datenbank. Abgerufen von <http://www.mysql.de> (Letzter Zugriff: Juni 2015)
- NATIONAL RESEARCH COUNCIL, REDISCOVERING GEOGRAPHY COMMITTEE (1997). *Rediscovering Geography: New Relevance for Science and Society*. Washington.
- NWB-TEAM. (2006). Network Workbench Tool. Abgerufen von <http://nwb.cns.iu.edu/> (Letzter Zugriff: Juni 2015)
- PFEFFER, J. (2013). txt2pajek - Software zur Umwandlung von Textfiles zu Pajekfiles. Abgerufen von <http://www.pfeffer.at/txt2pajek/> (Letzter Zugriff: Juni 2015)

- PFEFFER, J., MRVAR, A., & BATAGELJ, V. (2013). *txt2pajek: Creating Pajek Files from Text Files. Technical Report*.
- PORTER, M. F. (1980). An algorithm for suffix stripping. *Program: electronic library and information systems*, 14(3), S. 130–137.
- ROSCH, E. H. (1973). Natural Categories. *Cognitive Psychology*, 4(3), S. 328–350.
- SALVINI, M. M. (2012). Spatialization von nutzergenerierten Inhalten für die explorative Analyse des globalen Städteneetzes. Dissertation. Geographisches Institut der Universität Zürich.
- SKUPIN, A. (2000). From metaphor to method: cartographic perspectives on information visualization, in: *Proceedings of the IEEE Symposium on Information Visualization 2000*, S. 91–97.
- SKUPIN, A. (2002a). A cartographic approach to visualizing conference abstracts. *IEEE Computer Graphics and Applications*, 22(1), S. 50–58.
- SKUPIN, A. (2002b). On geometry and transformation in map-like information visualization, in: K. Börner & C. Chen (Hrsg.), *Visual Interfaces to Digital Libraries*, S. 161–170. Springer-Verlag Berlin Heidelberg.
- SKUPIN, A. (2004). The world of geography: Visualizing a knowledge domain with cartographic means. *Proceedings of the National Academy of Sciences of the United States of America*, 101(1), S. 5274–5278.
- SKUPIN, A. (2009). Discrete and continuous conceptualizations of science: Implications for knowledge domain visualization. *Journal of Informetrics*, 3(3), S. 233–245.
- SKUPIN, A. (2010). In Terms Of Geography, in: K. Börner (Hrsg.), *Atlas of Science*, S. 102–105. MIT Press.
- SKUPIN, A., & AGARWAL, P. (2008). Introduction: What is a Self-Organizing Map?, in: *Self-Organising Maps: Applications in Geographic Information Science*, S. 1–20. John Wiley & Sons, Ltd.
- SKUPIN, A., BIBERSTINE, J., & BÖRNER, K. (2013). Visualizing the topical structure of the medical sciences: a self-organizing map approach. *PloS one*, 8(3), S. e58779.
- SKUPIN, A., & BUTTENFIELD, B. P. (1996). Spatial metaphors for visualizing very large data archives, in: *Proceedings of GIS/LIS'96*, S. 607–617. Denver, CO.
- SKUPIN, A., & BUTTENFIELD, B. P. (1997). Spatial Metaphors for Visualizing Information Spaces, in: *Proceedings of ACSM/ASPRS Annual Convention and Exhibition*, S. 116–125. Seattle, WA.
- SKUPIN, A., & ESPERBÉ, A. (2008). Towards High-Resolution Self-Organizing Maps of Geographic Features, in: M. Dodge, M. McDerby, & M. Turner (Hrsg.), *Geographic Visualization: Concepts, Tools and Applications*, S. 159–181. John Wiley & Sons, Ltd.
- SKUPIN, A., & ESPERBÉ, A. (2011). An alternative map of the United States based on an n-dimensional model of geographic space. *Journal of Visual Languages & Computing*, 22(4), S. 290–304.

- SKUPIN, A., & FABRIKANT, S. I. (2003). Spatialization Methods: A Cartographic Research Agenda for Non-geographic Information Visualization. *Cartography and Geographic Information Science*, 30(2), S. 99–119.
- SKUPIN, A., & HAGELMAN, R. (2005). Visualizing demographic trajectories with self-organizing maps. *GeoInformatica*, 9(2), S. 159–179.
- SKUPIN, A., & DE JONGH, C. (2005). Visualizing the ICA – A Content-based Approach, in: *Proceedings of the 22nd International Cartographic Conference*. A Coruña, Spain.
- SLOCUM, T. A., MACMASTER, R. B., KESSLER, F. C., & HOWARD, H. H. (2009). *Thematic Cartography and Geovisualization, Third Edition*. Upper Saddle River (USA): Pearson.
- STEVENS, A., & COUPE, P. (1978). Distortions in judged spatial relations. *Cognitive psychology*, 10(4), S. 422–437.
- STEYVERS, M., & GRIFFITHS, T. (2007). Probabilistic topic models, in: T. K. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Hrsg.), *Handbook of Latent Semantic Analysis*. Erlbaum, Lawrence.
- TANG, J., MENG, Z., NGUYEN, X., MEI, Q., & ZHANG, M. (2014). Understanding the Limiting Factors of Topic Modeling via Posterior Contraction Analysis, in: *Proceedings of the 31st International Conference on Machine Learning*.
- TOBLER, W. R. (1970). A Computer Movie Simulation Urban Growth in the Detroit Region. *Economic Geography*, 46(2), S. 234–240.
- TORY, M., & MOLLER, T. (2004). Rethinking Visualization: A High-Level Taxonomy, in: *IEEE Symposium on Information Visualization*, S. 151–158.
- WISE, J. A., THOMAS, J. J., PENNOCK, K., LANTRIP, D., POTTIER, M., SCHUR, A., ET AL. (1995). Visualizing the Non-Visual: Spatial analysis and interaction with information from text documents, in: *Information Visualization, 1995. Proceedings.*, S. 51 – 58. Atlanta, GA.

## ANHANG

---

A	Liste ausgeschlossener Worte .....	75
B	Themenübersicht Topic Modeling .....	77
C	Fragebogen A.....	79
D	Auswertungen Excel.....	103
E	Auswertung SPSS .....	105
F	Persönliche Erklärung.....	117



## A LISTE AUSGESCHLOSSENER WÖRTE

Die Wörter, welche man in der Datenverarbeitung nicht berücksichtigen wollte, findet man in der Stoppwortliste online<sup>17</sup> Diese Liste ist Bestandteil des Natural Language Toolkit (Bird et al. 2009). In dieser Arbeit wird sie zum Ausschluss unerwünschter Wörter im Topic Modeling sowie für die Berechnung des TF/IDF-Wert genutzt.

A-C	D-H	I-O	P-T	U-Z
a	did	i	s	under
about	do	if	same	until
above	does	in	she	up
after	doing	into	should	very
again	don	is	so	was
against	down	it	some	we
all	during	its	such	were
am	each	itself	t	what
an	few	just	than	when
and	for	me	that	where
any	from	more	the	which
are	further	most	their	while
as	had	my	theirs	who
at	has	myself	them	whom
be	have	no	themselves	why
because	having	nor	then	will
been	he	not	there	with
before	her	now	these	you
being	here	of	they	your
below	hers	off	this	yours
between	herself	on	those	yourself
both	him	once	through	yourselves
but	himself	only	to	
by	his	or	too	
can	how	other		
		our		
		ours		
		ourselves		
		out		
		over		
		own		

<sup>17</sup> <http://www.nltk.org/book/ch02.html> (Letzter Zugriff: 21.04.2015)



## B THEMENÜBERSICHT TOPIC MODELING

TOPIC 1	TOPIC 2	TOPIC 3	TOPIC 4	TOPIC 5
cartography research use students experience knowledge education geography course training work teaching learning theory university studies children role understanding subject questions way develop study representation perception related school practice understand communication language present science subjects program applied courses field practical	process quality cartographic means problems processes aspects particular knowledge problem specific thus ie cartographic generalization cartographic products complex considered elements cartographic representation evaluation representation description applying defined purpose criteria cartographic design rules cartographic representations cartographic visualization appropriate modelling must requires many concepts discussed finally form relevant	geographical names history works names state work world historical geographical information first materials country russian geographical information system russian federation old also territory modern part period plans name political creation developed geographical preparation produced globe lot carried contribution documents historic drawn especially basis department geographical data	images topographic maps image accuracy used using surface satellite images terrain derived obtained scale contour lines relief dem accurate high quality satellite imagery technique aerial photographs l produced resolution earth topographic map detailed contours high resolution imagery dtm processing base size topography surfaces high features precise resulting high resolution satellite imagery	model method based generalization objects features methods process line algorithms set object applied parameters selection algorithm object oriented result original problem procedure proposed operations scale generalisation generalized shape function approach automatic rules automatically feature input according lines processing geometry constraints paper
TOPIC 6	TOPIC 7	TOPIC 8	TOPIC 9	TOPIC 10
data database spatial data spatial information geographic information integration management support data sets data model spatial databases existing information updating spatial analysis geographic information system geographic information systems available data base integrated data collection datasets format integrating spatial data infrastructure update layers spatial distribution data set data sources geographic information systems spatial databases dataset maintenance data processing attributes geographic data data quality geographic	maps map map design symbols used map use map production map series cartographers map projections map making good color series categories others various map reading map sheets map users map content map generalization cartographer legend many making map symbols symbolization art signs symbol elements one map projection first clear colours even today characteristics	analysis areas area using results study distribution population impact planning used local potential result related analyses case study quantpct evaluation methodology assessment procedures density boundaries studied case due along accessibility risk decision making taking account values calculation number range significant site located measures	time analysis space change city changes study de patterns studies research cities among phenomena evolution regarding thus urban areas china according methodology construction urban well using 2008 represents 2010 b space time growth 2001 built importance considering g time series comparison d 2006	users internet application services access available service based use tools geospatial data web metadata applications using software developed web based also provided web services distributed technologies user framework functions online standard functionality providing platform development provide open source resources open prototype implemented concept etc

TOPIC 11	TOPIC 12	TOPIC 13	TOPIC 14	TOPIC 15
design visualization research dynamic display methods users effective tools representation information use techniques visual user interactive task uncertainty navigation tasks useful animation real world used representing multimedia interaction electronic map real time view user interface tool usability paper characteristics communication paper describes representations include visual variables	information use context people location way like place understanding well decision making time specific often many environment interest places important individual ways also however issues relevant events within group appropriate especially mobile devices together groups adapted gps society critical current support place names	mapping use within needs need techniques field however requirements framework provide future date base demand work include including today needed improve variety availability challenge currently existing trends sources potential become limited required yet ensure major require increased source associated creating	environment region development complex conditions basis regions sustainable development nature monitoring natural resources territory changes etc processes russia management situation river resources problems landscape dynamics modeling estimation includes water investigation environmental influence natural landscapes ecological formation natural environment activity factors mapping structure realization	different approach used models based level structure results network given eg one proposed selected approaches set example relationship building solution therefore according using number properties complexity defined like road network levels type structures case study similar case terms value certain methods presented
TOPIC 16	TOPIC 17	TOPIC 18	TOPIC 19	TOPIC 20
new atlas well thematic maps form maps content created etc concept atlases world general published presentation forms called also example presented create creating product principles made special thematic cartography authors collection production terms thematic included necessary printed designed publication contents publishing including	project development production standards activities national countries implementation projects government sdi official country standard established involved infrastructure spain united states europe products common establishment part efforts report series south africa working organization national atlas plan current objectives cooperation european national park international public standardization	one described points although projection also first found part two fact applied show made another kind known possible transformation furthermore however could therefore number mentioned presented projections global respect due represent common even second usually shown coordinates parts whole describe	land use classification 3 d use based 2 area land cover 3 land remote sensing types following 4 interpretation 1 within type study soil shows units classes different vegetation 3 d visualization characteristics detailed rural areas especially important forest combined result effects combination landscape poland 3 d maps 5	system gis information technology systems development creation applications new digital developed application traditional developing introduced software including discussed various computer products information system producing gis technology processing digital mapping integrated digital cartography chart demands required developments digital maps paper making technologies digital map future gis applications gis based

## C FRAGEBOGEN A

---

Der komplette Fragebogen A ist in den nächsten Seiten dargestellt. Der Aufbau für Fragebogen B ist derselbe, nur sind einzelne Visualisierungen ausgetauscht. Vergleiche dafür das Kapitel 3.5.4. Zusätzlich muss man erwähnen, dass einzelne Formatierungsvorlagen von der OnlineUmfragen.com-Plattform hier nicht übernommen werden. Während %BLOCK...BLOCK% den Text auf der Plattform in ein Blocksatz umformatiert, fügt der Befehl %BILD ein Bild ein.

### Vorwort

#### Masterarbeit Oliver Deseö: Informationslandschaften im Vergleich I

Sie sind eingeladen, an einer empirischen Studie über die Wirkung von **Informationsvisualisierungen** teilzunehmen. Wir möchten dabei die Wirkung von Visualisierungstypen vergleichen, die eine **Informationslandschaft darstellen**. Diese Informationslandschaft visualisiert den Inhalt von Konferenzbeiträgen, die im Rahmen der internationalen Konferenzen der **Internationalen Kartographischen Vereinigung (IKV) von 1993-2013** publiziert wurden.

Diese Studie wird von Oliver Deseö im Zusammenhang mit seiner Masterarbeit durchgeführt. Die Masterarbeit wird von Prof. Dr. Sara Fabrikant vom Geographischen Institut der Universität Zürich geleitet.

#### Hinweis

Bitte verwenden Sie während der gesamten Studie **nicht** den «Zurück»-Button Ihres Browsers. Umso wichtiger ist es, dass Sie die Fragen sorgfältig durchlesen und beantworten. Beachten Sie auch, dass während der gesamten Studie die Zeit aufgezeichnet wird. Verweilen Sie also nur so lange wie nötig bei den einzelnen Aufgaben. Ausserdem sind während der gesamten Studie **keine Hilfsmittel** erlaubt.

#### Dauer

45 min

#### Autor

Oliver Deseö, Universität Zürich

### Persönliche Informationen

Bitte beantworten Sie zuerst ein paar Fragen zu Ihrem **Hintergrund**.

Was ist Ihr Geschlecht?

- weiblich  
 männlich

Wie alt sind Sie (in Jahren)?

Alter

Was ist der höchste Bildungsgrad, den Sie bisher erlangt haben?

- Matura/Abitur  
 Bachelordiplom  
 Masterdiplom  
 Doktorat  
 Professur  
 Anderes:

Befinden Sie sich im Geographie-Studium?

- Ja, im Hauptfach
- Ja, im Nebenfach.
- Nein, ich studiere etwas anderes.
- Nein, ich studiere gar nicht bzw. habe nicht studiert

In welcher Studienrichtung haben Sie den vorhin beschriebenen Bildungsgrad erlangt?

- Geographie
- Anderes:

Da Sie Geographie im Hauptfach studieren: In welche Fachrichtung spezialisieren Sie sich?

- Humangeographie
- Fernerkundung
- GIScience
- Physische Geographie
- Wirtschaftsgeographie
- Keine Spezialisierung

Da Sie einen Abschluss in Geographie haben: Welche Spezialisierung haben Sie gewählt?

- Physische Geographie
- Wirtschaftsgeographie
- Humangeographie
- Fernerkundung
- GIScience
- Keine Spezialisierung

Da Sie im Hauptfach etwas anderes studieren: Was studieren Sie im Hauptfach?

Wurde Ihnen von einer professionellen Fachperson (Optiker, Augenarzt) schon einmal gesagt, dass Sie eine Sehschwäche (z.B. Brille / Kontaktlinsen nötig, Farbenblindheit) haben?

- ja
- nein

Falls ja: Welche Art von Sehschwäche haben Sie (z.B. Brillenträger, Kontaktlinsen, Farbenblindheit)?

- Brillenträger oder Kontaktlinsen
- Farbenblindheit
- Anderes:

Wie gut kennen Sie sich in folgenden Bereichen aus?

	schlecht				gut
Karten lesen	<input type="radio"/>				
Karten selber gestalten	<input type="radio"/>				
Geographische Informationssysteme (GIS)	<input type="radio"/>				
Graphik-Design/Graphische Gestaltung	<input type="radio"/>				
Informationsvisualisierung gestalten	<input type="radio"/>				

### Probeaufgaben

%BLOCKSie sehen unten zwei Informationslandschaften, die Sie auch wieder im Hauptteil dieser Befragung antreffen werden.

%BILD%ueberblick\_klein.jpg%

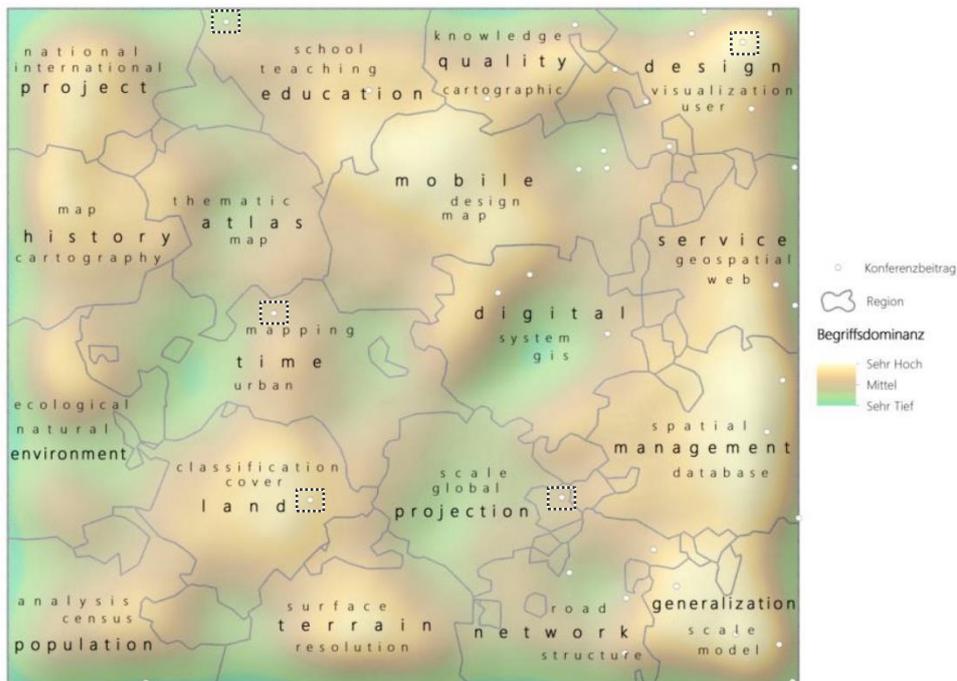
Die dargestellten **Regionen** stellen die Konferenzthemen dar, die sich aus dem Inhalt der Konferenzbeiträge der **International Kartographischen Vereinigung (IKV)** über die Jahre 1993-2013 ergeben. Die Regionen sind mit **drei Begriffen** beschriftet, die den Inhalt der Regionen aussagekräftig beschreiben. Einzelne Konferenzbeiträge sind als Punkte in der Informationslandschaft dargestellt. Als Zusatzinformation dient die sogenannte **Begriffsdominanz**, welche die Wichtigkeit der drei am häufigsten vorkommenden Begriffe im Verhältnis zu allen in der Informationslandschaft vorkommenden Begriffen beschreibt. Ist die Begriffsdominanz an einem Ort in der Landschaft hoch, bedeutet dies, dass die drei dort aufgeführten Begriffe in dieser Region, beziehungsweise in den einzelnen dort vorkommenden Konferenzbeiträgen überdurchschnittlich häufig vorkommen.

Sie werden nun ein paar Probeaufgaben lösen um das Konzept der Informationslandschaft näher kennenzulernen.

Bitte beantworten Sie die folgenden Fragen, indem Sie mit der Maus auf eines der 5 (fünf) vorgegebenen Rechtecke in der Informationslandschaft klicken. Ein **grünes Kreuz** erscheint. Möchten Sie Ihre Antwort ändern, dann klicken Sie zuerst auf das von Ihnen gesetzte grüne Kreuz um ihre Antwort rückgängig zu machen und wählen Sie danach eine andere Antwortmöglichkeit aus. Sie können erst zur nächsten Frage gelangen, wenn Sie die gestellte Aufgabe richtig gelöst haben.

BLOCK%

Wählen Sie den Beitrag mit der **höchsten Begriffsdominanz** aus den fünf in einem Quadrat eingerahmten Beiträgen, indem Sie auf **das jeweilige Quadrat** klicken.



Wählen Sie den Beitrag mit der **tiefsten Begriffsdominanz** aus den fünf in einem Quadrat eingerahmten Beiträgen, indem Sie auf **das jeweilige Quadrat** klicken.







Über welches Thema der IKV-Konferenz in Durban (2003) wurden die **meisten** Konferenzartikel geschrieben?  
 Hinweis: Es werden nur Beiträge der IKV-Konferenz in Durban im Jahre 2003 als weiße Punkte dargestellt. Beiträge aus anderen Jahren werden nicht dargestellt.

- design** | visualization | user
- national | international | **project**
- classification | cover | **land**
- mobile** | design | map
- school | teaching | **education**
- spatial | **management** | database

Begründen Sie Ihre Antwort (Stichworte genügen):

---

Wie sicher sind Sie sich bei Ihrer Antwort?

- sehr sicher
- sicher
- eher sicher
- eher unsicher
- unsicher
- sehr unsicher

Weitere Bemerkungen zu dieser Frage:

---



Über welches Thema der IKV-Konferenz in Stockholm (1997) wurden die **wenigsten** Konferenzartikel geschrieben?  
 Hinweis: Es werden nur Beiträge der IKV-Konferenz in Stockholm im Jahre 1997 als weiße Punkte dargestellt. Beiträge aus anderen Jahren werden nicht dargestellt.

- knowledge | **quality** | cartographic
- thematic | **atlas** | map
- scale | global | **projection**
- mapping | **time** | urban
- generalization** | scale | model
- mobile** | design | map

Begründen Sie Ihre Antwort (Stichworte genügen):

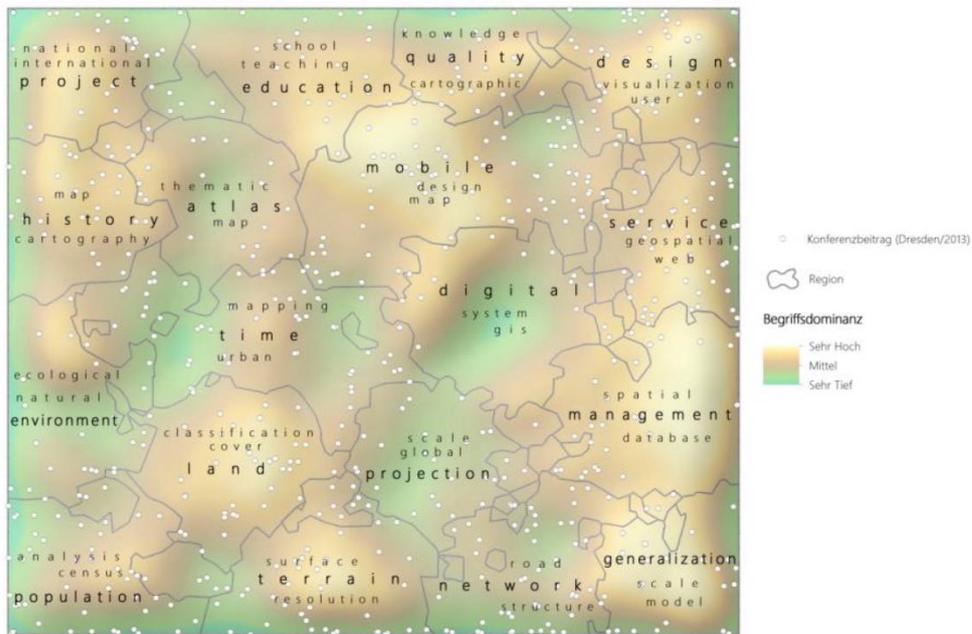
---

Wie sicher sind Sie sich bei Ihrer Antwort?

- sehr sicher
- sicher
- eher sicher
- eher unsicher
- unsicher
- sehr unsicher

Weitere Bemerkungen zu dieser Frage:

---



Über welches Thema der ICA-Konferenz in Dresden (2013) wurden die **meisten** Konferenzartikel geschrieben?  
 Hinweis: Es werden nur Beiträge der ICA-Konferenz in Dresden im Jahre 2013 als weiße Punkte dargestellt. Beiträge aus anderen Jahren werden nicht dargestellt.

- school | teaching | **education**
- surface | **terrain** | resolution
- analysis | census | **population**
- spatial | **management** | database
- mobile** | design | map
- generalization** | scale | model

Begründen Sie Ihre Antwort (Stichworte genügen):

Wie sicher sind Sie sich bei Ihrer Antwort?

- sehr sicher
- sicher
- eher sicher
- eher unsicher
- unsicher
- sehr unsicher

Weitere Bemerkungen zu dieser Frage:



Vergleichen Sie die inhaltliche Ähnlichkeit zwischen den Konferenzbeiträgen von **1** und **A** mit der inhaltlichen Ähnlichkeit zwischen **1** und **B**.

- 5 (**1** und **A** sind sich viel ähnlicher)
- 4
- 3
- 2
- 1 (**A** und **B** sind **1** gleich ähnlich)
- 2
- 3
- 4
- 5 (**1** und **B** sind sich viel ähnlicher)

Begründen Sie Ihre Antwort (Stichworte genügen):

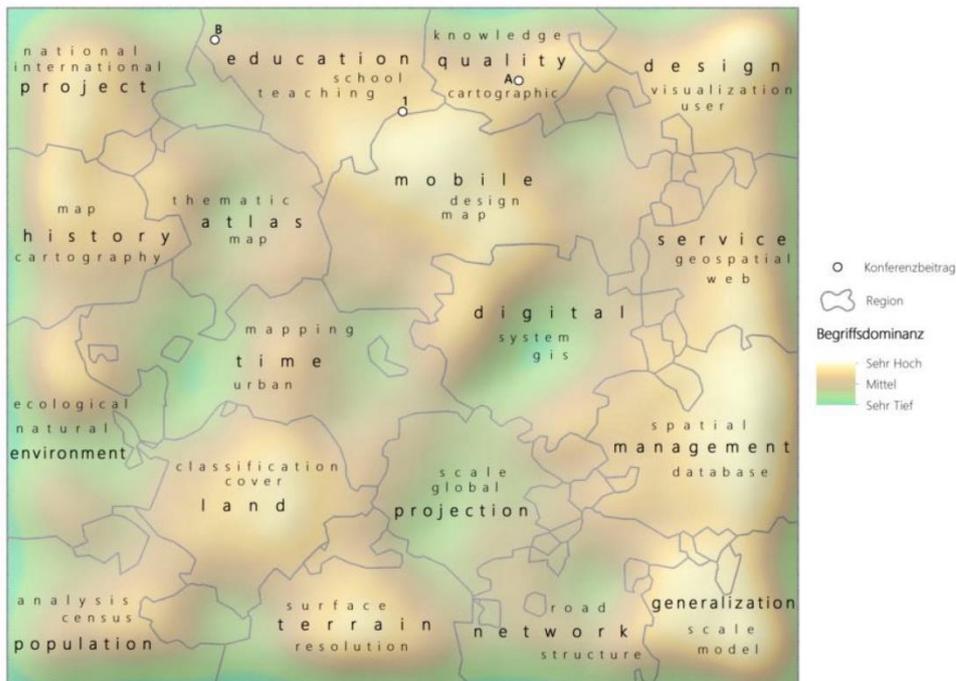
Welche Faktoren hatten einen Einfluss auf Ihre Entscheidung in der vorherigen Aufgabe?

	keinen Einfluss	wenig Einfluss	mittlerer Einfluss	viel Einfluss
Regionsgrenzen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Regionszugehörigkeit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Distanz zwischen den Konferenzbeiträgen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Begriffsdominanz	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Oberflächenbeschaffenheit der Landschaft	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Wie sicher sind Sie sich bei Ihrer Antwort?

- sehr sicher
- sicher
- eher sicher
- eher unsicher
- unsicher
- sehr unsicher

Weitere Bemerkungen zu dieser Frage:



Vergleichen Sie die inhaltliche Ähnlichkeit zwischen den Konferenzbeiträgen von **1** und **A** mit der inhaltlichen Ähnlichkeit zwischen **1** und **B**.

- 5 (**1** und **A** sind sich viel ähnlicher)
- 4
- 3
- 2
- 1 (**A** und **B** sind **1** gleich ähnlich)
- 2
- 3
- 4
- 5 (**1** und **B** sind sich viel ähnlicher)

Begründen Sie Ihre Antwort (Stichworte genügen):

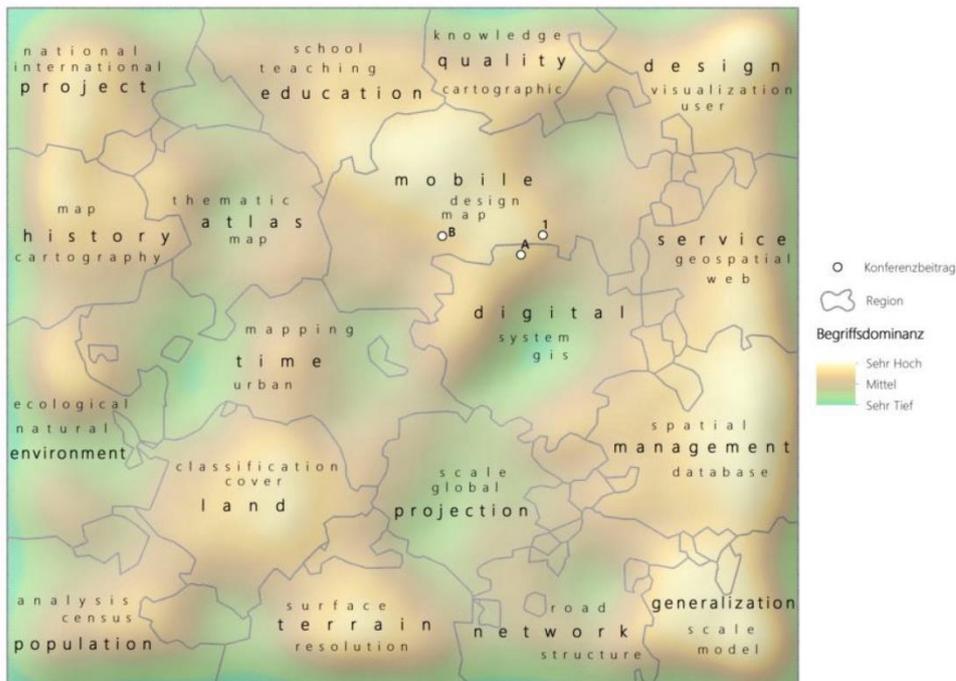
Welche Faktoren hatten einen Einfluss auf Ihre Entscheidung in der vorherigen Aufgabe?

	keinen Einfluss	wenig Einfluss	mittlerer Einfluss	viel Einfluss
Begriffsdominanz	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Oberflächenbeschaffenheit der Landschaft	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Distanz zwischen den Konferenzbeiträgen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Regionszugehörigkeit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Regionsgrenzen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Wie sicher sind Sie sich bei Ihrer Antwort?

- sehr sicher
- sicher
- eher sicher
- eher unsicher
- unsicher
- sehr unsicher

Weitere Bemerkungen zu dieser Frage:



Vergleichen Sie die inhaltliche Ähnlichkeit zwischen den Konferenzbeiträgen von **1** und **A** mit der inhaltlichen Ähnlichkeit zwischen **1** und **B**.

- 5 (**1** und **A** sind sich viel ähnlicher)
- 4
- 3
- 2
- 1 (**A** und **B** sind **1** gleich ähnlich)
- 2
- 3
- 4
- 5 (**1** und **B** sind sich viel ähnlicher)

Begründen Sie Ihre Antwort (Stichworte genügen):

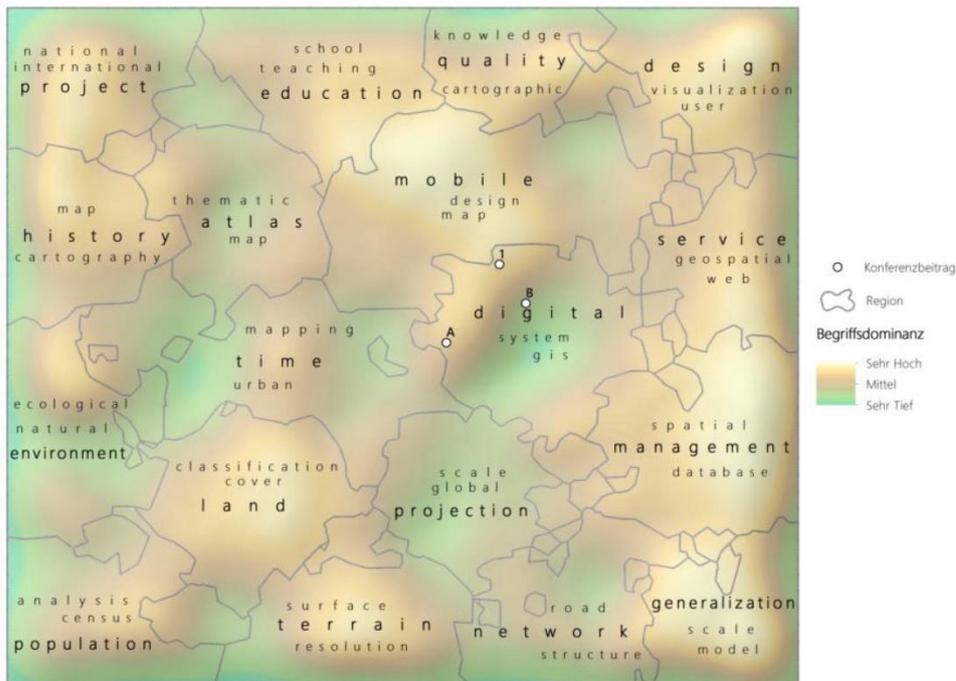
Welche Faktoren hatten einen Einfluss auf Ihre Entscheidung in der vorherigen Aufgabe?

	keinen Einfluss	wenig Einfluss	mittlerer Einfluss	viel Einfluss
Distanz zwischen den Konferenzbeiträgen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Regionsgrenzen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Regionszugehörigkeit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Oberflächenbeschaffenheit der Landschaft	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Begriffsdominanz	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Wie sicher sind Sie sich bei Ihrer Antwort?

- sehr sicher
- sicher
- eher sicher
- eher unsicher
- unsicher
- sehr unsicher

Weitere Bemerkungen zu dieser Frage:



Vergleichen Sie die inhaltliche Ähnlichkeit zwischen den Konferenzbeiträgen von **1** und **A** mit der inhaltlichen Ähnlichkeit zwischen **1** und **B**.

- 5 (**1** und **A** sind sich viel ähnlicher)
- 4
- 3
- 2
- 1 (**A** und **B** sind **1** gleich ähnlich)
- 2
- 3
- 4
- 5 (**1** und **B** sind sich viel ähnlicher)

Begründen Sie Ihre Antwort (Stichworte genügen):

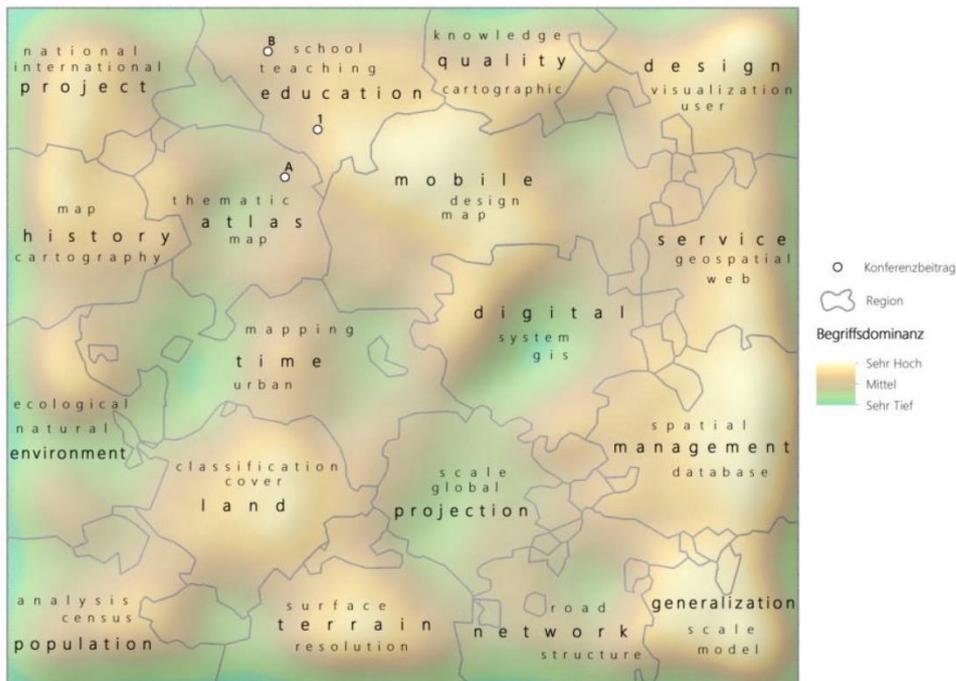
Welche Faktoren hatten einen Einfluss auf Ihre Entscheidung in der vorherigen Aufgabe?

	keinen Einfluss	wenig Einfluss	mittlerer Einfluss	viel Einfluss
Regionsgrenzen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Oberflächenbeschaffenheit der Landschaft	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Distanz zwischen den Konferenzbeiträgen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Begriffsdominanz	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Regionszugehörigkeit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Wie sicher sind Sie sich bei Ihrer Antwort?

- sehr sicher
- sicher
- eher sicher
- eher unsicher
- unsicher
- sehr unsicher

Weitere Bemerkungen zu dieser Frage:



Vergleichen Sie die inhaltliche Ähnlichkeit zwischen den Konferenzbeiträgen von **1** und **A** mit der inhaltlichen Ähnlichkeit zwischen **1** und **B**.

- 5 (**1** und **A** sind sich viel ähnlicher)
- 4
- 3
- 2
- 1 (**A** und **B** sind **1** gleich ähnlich)
- 2
- 3
- 4
- 5 (**1** und **B** sind sich viel ähnlicher)

Begründen Sie Ihre Antwort (Stichworte genügen):

Welche Faktoren hatten einen Einfluss auf Ihre Entscheidung in der vorherigen Aufgabe?

	keinen Einfluss	wenig Einfluss	mittlerer Einfluss	viel Einfluss
Distanz zwischen den Konferenzbeiträgen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Regionszugehörigkeit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Regionsgrenzen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Oberflächenbeschaffenheit der Landschaft	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Begriffsdominanz	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Wie sicher sind Sie sich bei Ihrer Antwort?

- sehr sicher
- sicher
- eher sicher
- eher unsicher
- unsicher
- sehr unsicher

Weitere Bemerkungen zu dieser Frage:



Vergleichen Sie die inhaltliche Ähnlichkeit zwischen den Konferenzbeiträgen von **1** und **A** mit der inhaltlichen Ähnlichkeit zwischen **1** und **B**.

- 5 (**1** und **A** sind sich viel ähnlicher)
- 4
- 3
- 2
- 1 (**A** und **B** sind **1** gleich ähnlich)
- 2
- 3
- 4
- 5 (**1** und **B** sind sich viel ähnlicher)

Begründen Sie Ihre Antwort (Stichworte genügen):

Welche Faktoren hatten einen Einfluss auf Ihre Entscheidung in der vorherigen Aufgabe?

	keinen Einfluss	wenig Einfluss	mittlerer Einfluss	viel Einfluss
Begriffsdominanz	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Distanz zwischen den Konferenzbeiträgen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Oberflächenbeschaffenheit der Landschaft	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Regionsgrenzen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Regionszugehörigkeit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Wie sicher sind Sie sich bei Ihrer Antwort?

- sehr sicher
- sicher
- eher sicher
- eher unsicher
- unsicher
- sehr unsicher

Weitere Bemerkungen zu dieser Frage:



Vergleichen Sie die inhaltliche Ähnlichkeit zwischen den Konferenzbeiträgen von **1** und **A** mit der inhaltlichen Ähnlichkeit zwischen **1** und **B**.

- 5 (**1** und **A** sind sich viel ähnlicher)
- 4
- 3
- 2
- 1 (**A** und **B** sind **1** gleich ähnlich)
- 2
- 3
- 4
- 5 (**1** und **B** sind sich viel ähnlicher)

Begründen Sie Ihre Antwort (Stichworte genügen):

Welche Faktoren hatten einen Einfluss auf Ihre Entscheidung in der vorherigen Aufgabe?

	keinen Einfluss	wenig Einfluss	mittlerer Einfluss	viel Einfluss
Regionsgrenzen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Oberflächenbeschaffenheit der Landschaft	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Begriffsdominanz	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Distanz zwischen den Konferenzbeiträgen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Regionzugehörigkeit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Wie sicher sind Sie sich bei Ihrer Antwort?

- sehr sicher
- sicher
- eher sicher
- eher unsicher
- unsicher
- sehr unsicher

Weitere Bemerkungen zu dieser Frage:



Vergleichen Sie die inhaltliche Ähnlichkeit zwischen den Konferenzbeiträgen von **1** und **A** mit der inhaltlichen Ähnlichkeit zwischen **1** und **B**.

- 5 (**1** und **A** sind sich viel ähnlicher)
- 4
- 3
- 2
- 1 (**A** und **B** sind **1** gleich ähnlich)
- 2
- 3
- 4
- 5 (**1** und **B** sind sich viel ähnlicher)

Begründen Sie Ihre Antwort (Stichworte genügen):

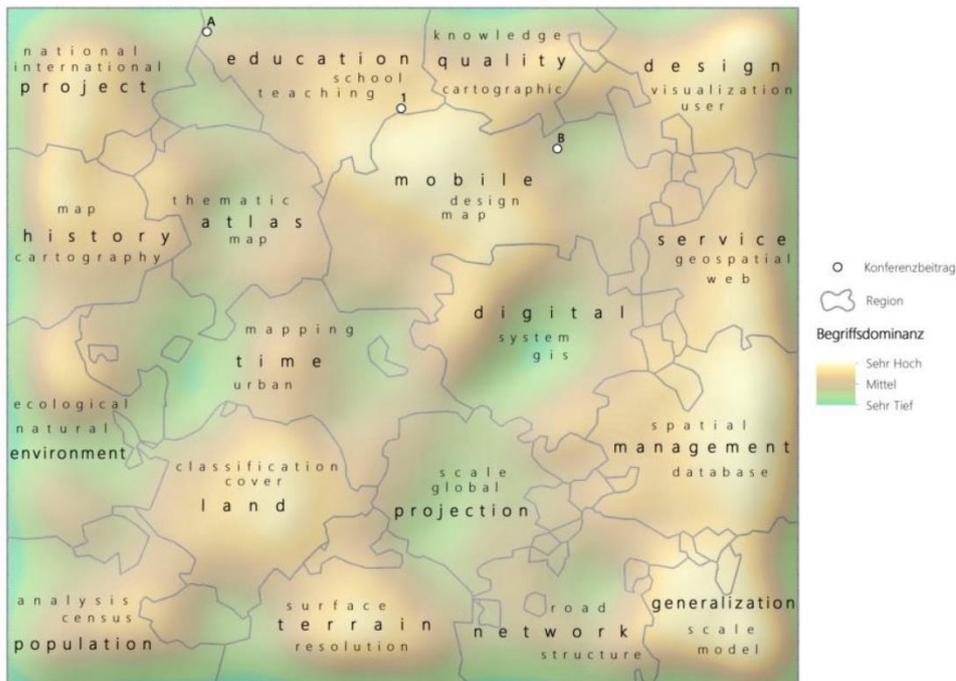
Welche Faktoren hatten einen Einfluss auf Ihre Entscheidung in der vorherigen Aufgabe?

	keinen Einfluss	wenig Einfluss	mittlerer Einfluss	viel Einfluss
Regionszugehörigkeit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Regionsgrenzen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Oberflächenbeschaffenheit der Landschaft	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Begriffsdominanz	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Distanz zwischen den Konferenzbeiträgen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Wie sicher sind Sie sich bei Ihrer Antwort?

- sehr sicher
- sicher
- eher sicher
- eher unsicher
- unsicher
- sehr unsicher

Weitere Bemerkungen zu dieser Frage:



Vergleichen Sie die inhaltliche Ähnlichkeit zwischen den Konferenzbeiträgen von **1** und **A** mit der inhaltlichen Ähnlichkeit zwischen **1** und **B**.

- 5 (**1** und **A** sind sich viel ähnlicher)
- 4
- 3
- 2
- 1 (**A** und **B** sind **1** gleich ähnlich)
- 2
- 3
- 4
- 5 (**1** und **B** sind sich viel ähnlicher)

Begründen Sie Ihre Antwort (Stichworte genügen):

Welche Faktoren hatten einen Einfluss auf Ihre Entscheidung in der vorherigen Aufgabe?

	keinen Einfluss	wenig Einfluss	mittlerer Einfluss	viel Einfluss
Oberflächenbeschaffenheit der Landschaft	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Distanz zwischen den Konferenzbeiträgen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Begriffsdominanz	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Regionszugehörigkeit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Regionsgrenzen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Wie sicher sind Sie sich bei Ihrer Antwort?

- sehr sicher
- sicher
- eher sicher
- eher unsicher
- unsicher
- sehr unsicher

Weitere Bemerkungen zu dieser Frage:





Vergleichen Sie die inhaltliche Ähnlichkeit zwischen den Konferenzbeiträgen von **1** und **A** mit der inhaltlichen Ähnlichkeit zwischen **1** und **B**.

- 5 (**1** und **A** sind sich viel ähnlicher)
- 4
- 3
- 2
- 1 (**A** und **B** sind **1** gleich ähnlich)
- 2
- 3
- 4
- 5 (**1** und **B** sind sich viel ähnlicher)

Begründen Sie Ihre Antwort (Stichworte genügen):

Welche Faktoren hatten einen Einfluss auf Ihre Entscheidung in der vorherigen Aufgabe?

	keinen Einfluss	wenig Einfluss	mittlerer Einfluss	viel Einfluss
Distanz zwischen den Konferenzbeiträgen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Regionsgrenzen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Regionszugehörigkeit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Begriffsdominanz	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Oberflächenbeschaffenheit der Landschaft	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Wie sicher sind Sie sich bei Ihrer Antwort?

- sehr sicher
- sicher
- eher sicher
- eher unsicher
- unsicher
- sehr unsicher

Weitere Bemerkungen zu dieser Frage:



Vergleichen Sie die inhaltliche Ähnlichkeit zwischen den Konferenzbeiträgen von **1** und **A** mit der inhaltlichen Ähnlichkeit zwischen **1** und **B**.

- 5 (**1** und **A** sind sich viel ähnlicher)
- 4
- 3
- 2
- 1 (**A** und **B** sind **1** gleich ähnlich)
- 2
- 3
- 4
- 5 (**1** und **B** sind sich viel ähnlicher)

Begründen Sie Ihre Antwort (Stichworte genügen):

Welche Faktoren hatten einen Einfluss auf Ihre Entscheidung in der vorherigen Aufgabe?

	keinen Einfluss	wenig Einfluss	mittlerer Einfluss	viel Einfluss
Oberflächenbeschaffenheit der Landschaft	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Begriffsdominanz	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Distanz zwischen den Konferenzbeiträgen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Regionszugehörigkeit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Regionsgrenzen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Wie sicher sind Sie sich bei Ihrer Antwort?

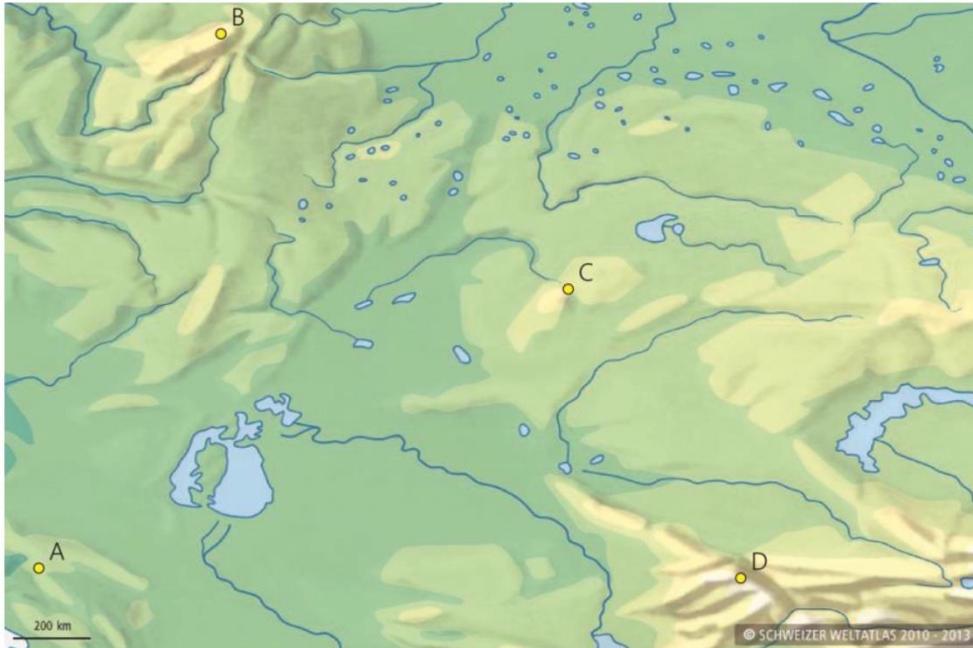
- sehr sicher
- sicher
- eher sicher
- eher unsicher
- unsicher
- sehr unsicher

Weitere Bemerkungen zu dieser Frage:

## Ende des Hauptteils

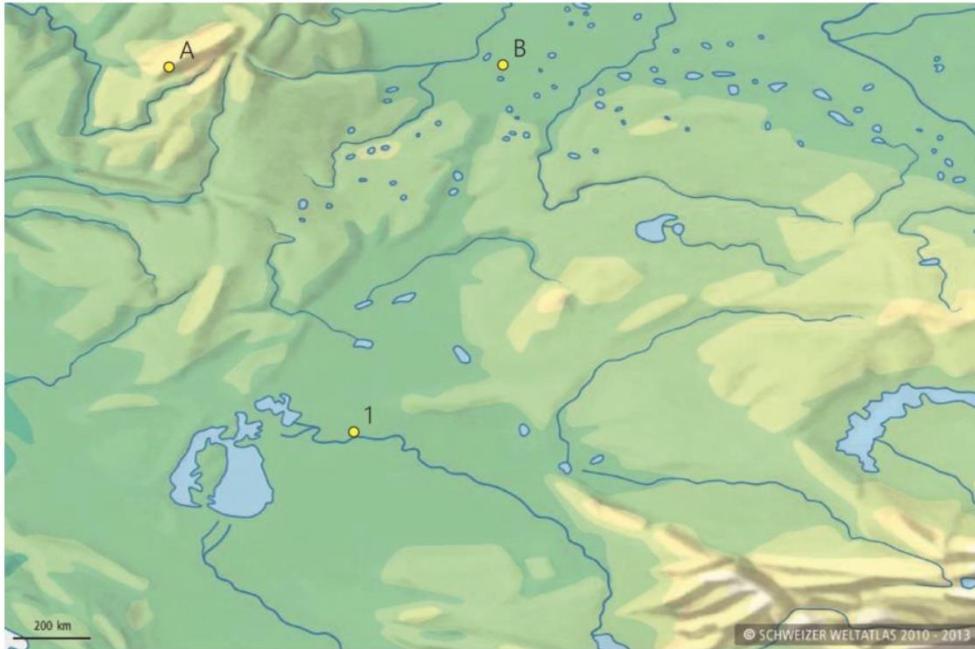
%BLOCKGleich haben Sie es geschafft.

Im letzten Teil der Studie möchten wir gerne mehr über Ihre Kartenkenntnisse wissen sowie Ihre Meinung zu den vorgestellten Visualisierungsarten erfahren. BLOCK%



Welches ist der höchste Punkt in dieser Landschaft?

- Punkt A
- Punkt B
- Punkt C
- Punkt D

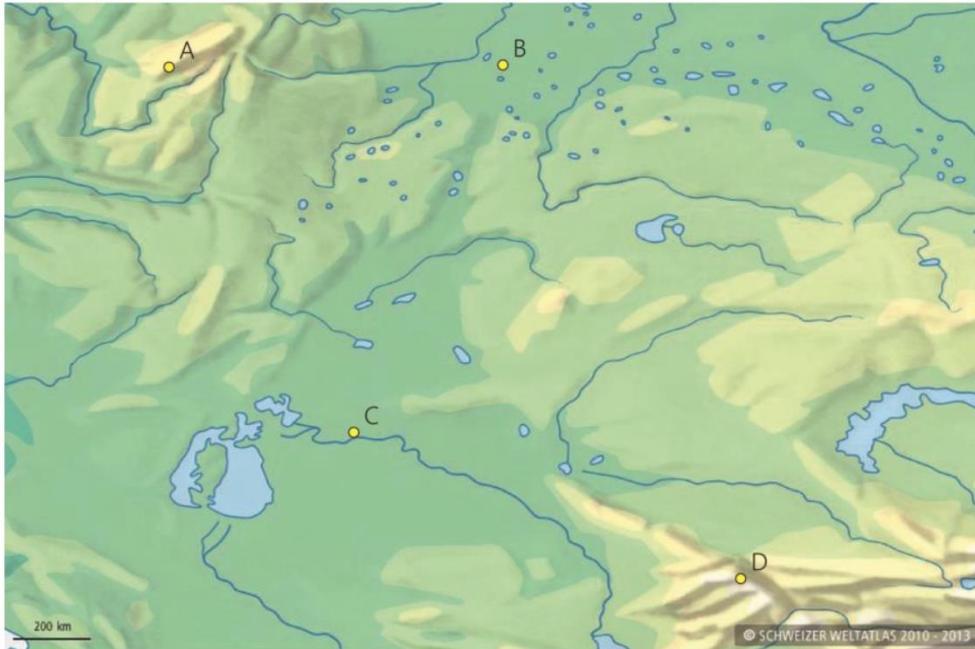


Vergleichen Sie die Ähnlichkeit zwischen den Punkten **1** und **A** mit der Ähnlichkeit zwischen **1** und **B**.

- 5 (**1** und **A** sind sich viel ähnlicher)
- 4
- 3
- 2
- 1 (**A** und **B** sind etwa gleich weit entfernt zu **1**)
- 2
- 3
- 4
- 5 (**1** und **B** sind sich viel ähnlicher)

Begründen Sie Ihre Antwort:

---



Rangieren Sie die folgenden Strecken nach Distanz. **Je grösser** die Distanz, **desto höher** die Strecke in der Rangliste.

Erstellen Sie bitte eine Rangliste mit 4 Elementen.

Hier nehmen!

von Punkt C nach Punkt A

von Punkt C nach Punkt D

von Punkt C nach Punkt B

von Punkt A nach Punkt B

Ihre Rangliste:

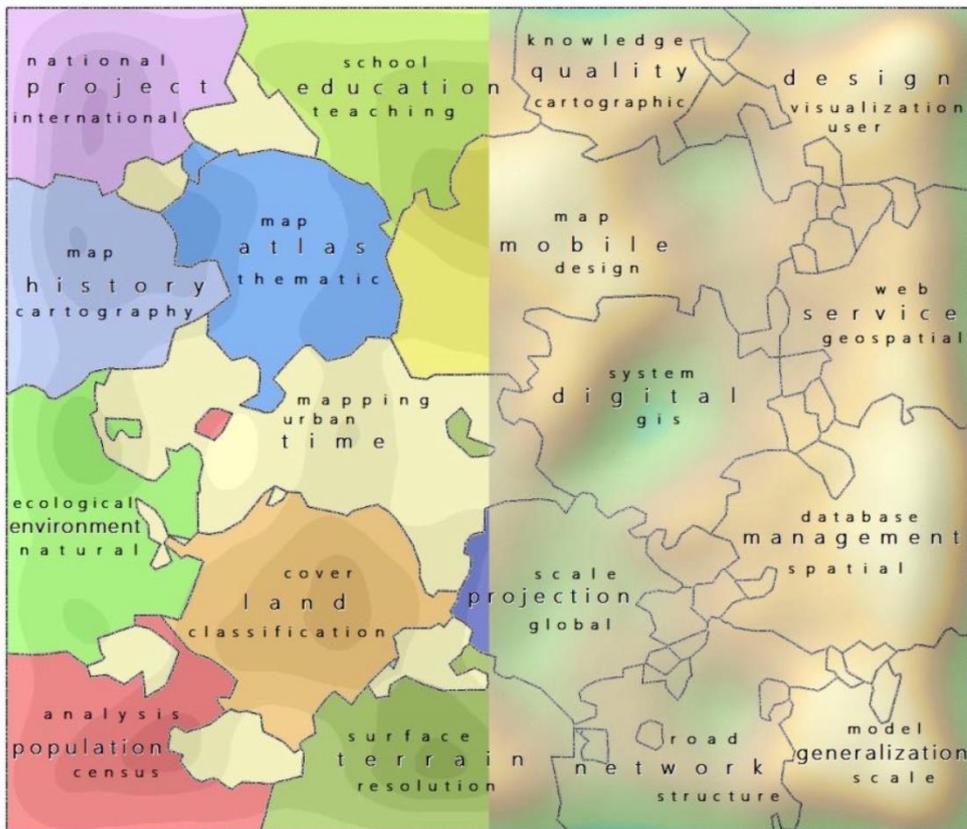
1.

2.

3.

(...)

Weitere Bemerkungen zu dieser Frage:



Bewerten Sie die Visualisierungstypen **A (links)** sowie **B (rechts)**.

	1 (Ziehe <b>Typ A</b> am meisten vor)	2	3	4	5 (Ziehe <b>Typ B</b> am meisten vor)
Bei welchen Fragetypen würden Sie Visualisierungstyp A oder B bevorzugen ?					
<b>Fragetyp:</b> Finden der höchsten/tiefsten Begriffsdominanz	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<b>Fragetyp:</b> Identifizieren der Region mit den meisten/wenigsten Konferenzbeiträgen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<b>Fragetyp:</b> Vergleichen der inhaltlichen Ähnlichkeit von Konferenzbeiträgen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<b>Insgesamt,</b>					
Welchen Visualisierungstyp A oder B würden Sie zur Beantwortung <b>aller Fragen</b> bevorzugen?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Weitere Bemerkungen zu dieser Frage:

Wenn Sie über die Ergebnisse der Studie auf dem Laufenden gehalten werden möchten, bitten wir Sie, nun ihre E-Mail-Adresse zu hinterlassen. Eine Kopie einer zukünftigen Publikation bzw. zukünftiger Publikationen wird Ihnen daraufhin zugestellt.

Bitte vervollständigen Sie die folgenden Angaben:

Vorname

Name

E-Mail

und ich sage... **Danke!**

Masterarbeit Oliver Deseö

Herzlichen Dank für Ihre Antworten!

Sie haben uns wesentlich geholfen, die Visualisierung von geographischen Metaphern in der Informationsvisualisierung besser zu verstehen.

**Autor**

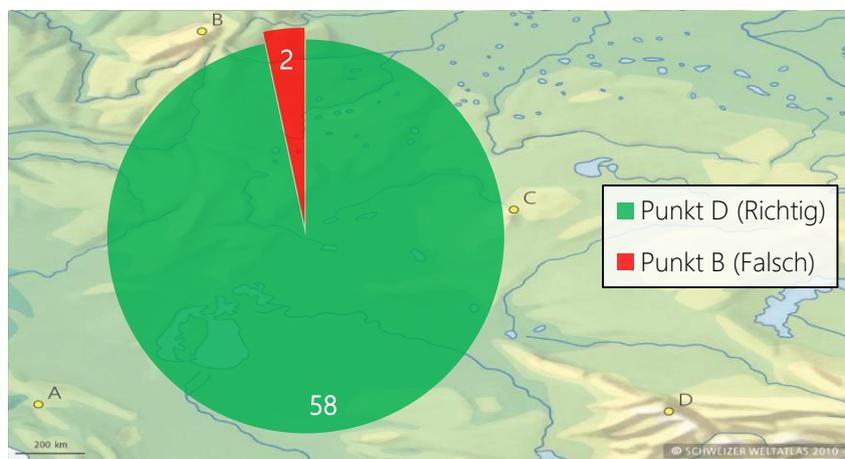
Oliver Deseö, Universität Zürich

## D AUSWERTUNGEN EXCEL

In diesem Teil des Anhangs sind alle Resultate der statistisch nicht-beantworteten Fragen der Studie. Dabei handelt es sich vor allem um die Kontrollfragen, die in dieser Arbeit nicht statistisch ausgewertet und diskutiert wurden. Dieses Kapitel soll die Resultate vervollständigen und weitere mögliche Studienideen bezüglich Informationslandschaften mit dieser Auswertung liefern

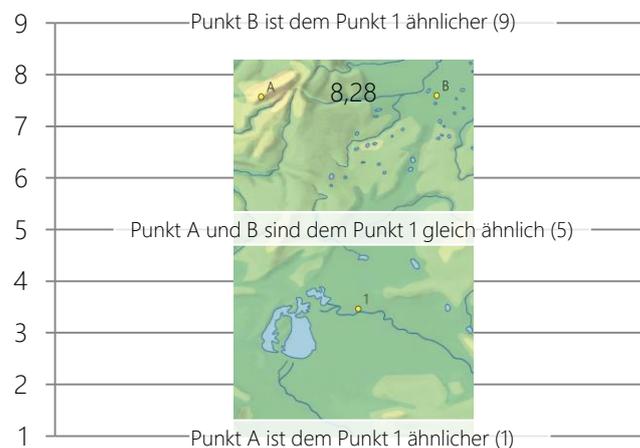
### D1 Kontrollfrage: Finden des höchsten Punkts

In dieser Aufgabe mussten die Probandinnen und Probanden den höchsten Punkt in einem Ausschnitt der europäischen Landschaft bestimmen. Sie konnten zwischen den vier Punkten A, B, C und D wählen, wobei der Punkt D am höchsten gelegen ist. Es handelt sich dabei um den Berg *Gora Manas*, der 4482m ü.M. hoch ist. Der Punkt B repräsentiert den Berg *Jamantau* mit 1640m ü.M. als zweithöchster Punkt dieser Auswahl.



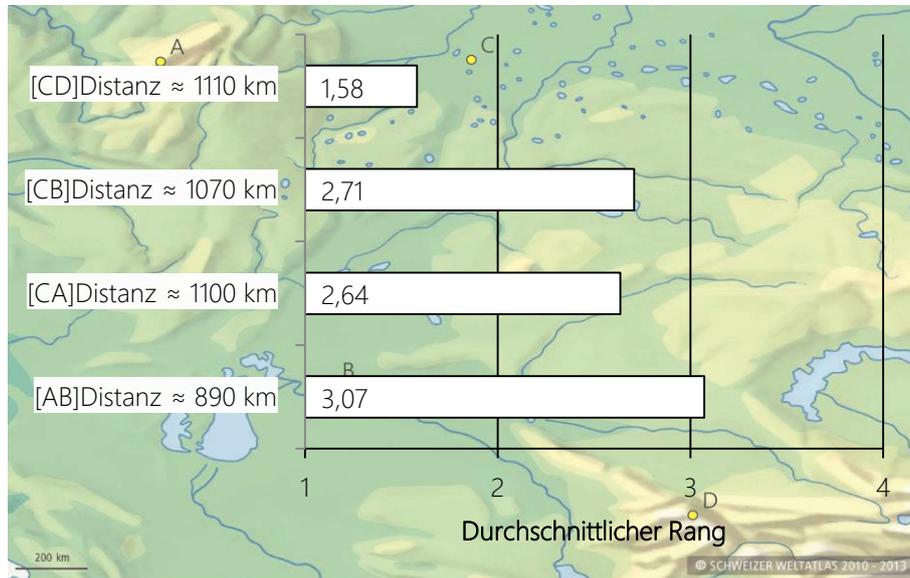
### D2 Kontrollfrage: Ähnlichkeitswahrnehmung

Analog zu den Ähnlichkeitsaufgaben in den Informationslandschaften mussten die Testpersonen einen Ausschnitt der europäischen Landschaft bearbeiten. Dabei sind die Distanzen von 1 zu A ungefähr 1100km und von 1 zu B 1070km. Der Durchschnitt von 8,28 bei einer Skala von 1 bis 9 weist auf ein deutliches Resultat hin. Dies kann mit dem gleichen Topographielevel und der etwas kleineren Distanz begründet werden. Diese Aufgabe wurde nicht in der Arbeit verwendet, da die geographische Metapher der Region in dieser Darstellung gefehlt hat und somit die Vergleichbarkeit mit den meisten Aufgaben nicht gegeben ist.



### D3 Kontrollfrage: Rangierung der Distanzen

Bei dieser Aufgabenstellung mussten die Teilnehmerinnen und Teilnehmer vier Strecken in einem Ausschnitt einer europäischen Reliefkarte anhand der Distanz rangieren. Dabei soll die längste Strecke den höchsten Rang (1) erhalten.



## E AUSWERTUNG SPSS

### E1 Nicht-parametrische Korrelationsanalyse nach Spearman: Einschätzung der Erfahrung – Geographie-Studium

Diese statistische Auswertung überprüft, ob die selbsteingeschätzte durchschnittliche Erfahrung in Karten lesen, Karten selber gestalten, Geographischen Informationssysteme (GIS), Graphik-Design/Graphische Gestaltung und Informationsvisualisierungen gestalten vom Geographie-Studium beeinflusst wird. In der deskriptiven Statistik ist zusätzlich die Attribuierung der Kontrollvariable «Ausbildung» erkennbar. Dabei erhalten alle Testpersonen, die im Hauptfach Geographie studieren oder studiert haben, einen Wert von «1». Die restlichen Ausbildungen werden mit einem Wert «2» statistisch verarbeitet. Aus dem Resultat erkennt man signifikante Korrelation der Ausbildung und der durchschnittlichen Erfahrung.

Deskriptive Statistik<sup>a</sup>

	N	Minimum	Maximum	Mittelwert	Standardabweichung
Ausbildung	34	1.0	1.0	1.000	.0000
Durchschnittliche Erfahrung	34	2.0	5.0	3.447	.7085
Gültige Werte (Listenweise)	34				

a. Ausbildung = Geographie-Hauptfach

Deskriptive Statistik<sup>a</sup>

	N	Minimum	Maximum	Mittelwert	Standardabweichung
Ausbildung	26	2.0	2.0	2.000	.0000
Durchschnittliche Erfahrung	26	1.4	4.2	2.608	.8717
Gültige Werte (Listenweise)	26				

a. Ausbildung = Andere

Korrelationen

		Ausbildung		Durchschnittliche Erfahrung
Spearman-Rho	Ausbildung	Korrelationskoeffizient	1.000	-.454**
		Sig. (2-seitig)	.	.000
		N	60	60
	Durchschnittliche Erfahrung	Korrelationskoeffizient	-.454**	1.000
		Sig. (2-seitig)	.000	.
		N	60	60

\*\* . Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

## E2 «GLM repeated measures»: Korrektheit bei Übersichtsaufgaben – Geographie-Studium

Deskriptive Statistiken					Innersubjektffaktoren				
					Maß: Korrektheit				
					Visualisierungstyp		Abhängige Variable		
					1	ÜbersichtTypD			
					2	ÜbersichtTypK			
					Zwischensubjektffaktoren				
					Wertelabel		N		
					Ausbildung	1	Geographie-Hauptfach	34	
						2	Andere	26	
Visualisierungstyp D (Übersicht)	Ausbildung	Mittelwert	Standardabweichung	N					
	Geographie-Hauptfach	1.618	.6520	34					
	Andere	1.462	.6469	26					
Gesamt		1.550	.6490	60					
Visualisierungstyp K (Übersicht)	Geographie-Hauptfach	1.529	.6622	34					
	Andere	1.423	.8086	26					
	Gesamt	1.483	.7247	60					

Maß: Korrektheit

Quelle		Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat	Nichtzentralitäts-Parameter	Beobachtete Schärfe <sup>a</sup>
Visualisierungstyp	Sphärizität angenommen	.118	1	.118	.534	.468	.009	.534	.111
	Greenhouse-Geisser	.118	1.000	.118	.534	.468	.009	.534	.111
	Huynh-Feldt	.118	1.000	.118	.534	.468	.009	.534	.111
	Untergrenze	.118	1.000	.118	.534	.468	.009	.534	.111
Visualisierungstyp * Ausbildung	Sphärizität angenommen	.018	1	.018	.082	.775	.001	.082	.059
	Greenhouse-Geisser	.018	1.000	.018	.082	.775	.001	.082	.059
	Huynh-Feldt	.018	1.000	.018	.082	.775	.001	.082	.059
	Untergrenze	.018	1.000	.018	.082	.775	.001	.082	.059
Fehler(Visualisierungstyp)	Sphärizität angenommen	12.848	58	.222					
	Greenhouse-Geisser	12.848	58.000	.222					
	Huynh-Feldt	12.848	58.000	.222					
	Untergrenze	12.848	58.000	.222					

a. Unter Verwendung von Alpha = .05 berechnet

Tests der Zwischensubjektffekte

Quelle		Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat	Nichtzentralitäts-Parameter	Beobachtete Schärfe <sup>a</sup>
Konstanter Term		134.004	1	134.004	366.102	.000	.863	366.102	1.000
Ausbildung		.254	1	.254	.693	.409	.012	.693	.130
Fehler		21.230	58	.366					

a. Unter Verwendung von Alpha = .05 berechnet

## E3 «GLM repeated measures»: Sicherheit bei Übersichtsaufgaben – Geographie-Studium

Deskriptive Statistiken					Innersubjektffaktoren				
					Maß: Sicherheit				
					Visualisierungstyp		Abhängige Variable		
					1	ÜberSicherTypD			
					2	ÜberSicherTypK			
					Zwischensubjektffaktoren				
					Wertelabel		N		
					Ausbildung	1	Geographie-Hauptfach	34	
						2	Andere	26	
Übersicht Sicherheit Typ D	Ausbildung	Mittelwert	Standardabweichung	N					
	Geographie-Hauptfach	2.794	.9933	34					
	Andere	3.596	1.1227	26					
Gesamt		3.142	1.1165	60					
Übersicht Sicherheit Typ K	Geographie-Hauptfach	2.765	1.0168	34					
	Andere	3.500	1.2490	26					
	Gesamt	3.083	1.1723	60					

Maß: Sicherheit

Quelle		Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat	Nichtzentralitäts-Parameter	Beobachtete Schärfe <sup>a</sup>
Visualisierungstyp	Sphärizität angenommen	.116	1	.116	.509	.479	.009	.509	.108
	Greenhouse-Geisser	.116	1.000	.116	.509	.479	.009	.509	.108
	Huynh-Feldt	.116	1.000	.116	.509	.479	.009	.509	.108
	Untergrenze	.116	1.000	.116	.509	.479	.009	.509	.108
Visualisierungstyp * Ausbildung	Sphärizität angenommen	.033	1	.033	.144	.706	.002	.144	.066
	Greenhouse-Geisser	.033	1.000	.033	.144	.706	.002	.144	.066
	Huynh-Feldt	.033	1.000	.033	.144	.706	.002	.144	.066
	Untergrenze	.033	1.000	.033	.144	.706	.002	.144	.066
Fehler(Visualisierungstyp)	Sphärizität angenommen	13.240	58	.228					
	Greenhouse-Geisser	13.240	58.000	.228					
	Huynh-Feldt	13.240	58.000	.228					
	Untergrenze	13.240	58.000	.228					

a. Unter Verwendung von Alpha = .05 berechnet

Tests der Zwischensubjektffekte

Quelle		Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat	Nichtzentralitäts-Parameter	Beobachtete Schärfe <sup>a</sup>
Konstanter Term		589.880	1	589.880	552.064	.000	.905	552.064	1.000
Ausbildung		8.705	1	8.705	8.147	.006	.123	8.147	.801
Fehler		61.973	58	1.068					

a. Unter Verwendung von Alpha = .05 berechnet

Bei dieser Auswertung ist zu beachten, dass ursprünglich der Wert 1 zur Einschätzung «Sehr sicher» zugeordnet wurde. Dies wurde für die Diagramme im Kapitel Resultate umgewandelt. Doch hier im Anhang gilt: je tiefer der Sicherheitswert, desto sicherer fühlten sich die Testpersonen.

### E4 <GLM repeated measures>: Korrektheit bei Ähnlichkeitsaufgaben – Geographie-Studium

Deskriptive Statistiken					Innersubjektfaktoren				
					Maß: Korrektheit				
					Visualisierungstyp Abhängige Variable				
					1 ÄhnlichKorrektTy pD				
					2 ÄhnlichKorrektTy pK				
					Zwischensubjektfaktoren				
					Wertelabel N				
					Ausbildung 1 Geographie-Hauptfach 34				
					2 Andere 26				
	Ausbildung	Mittelwert	Standardabweichung	N					
Korretheit Typ D	Geographie-Hauptfach	1,05882352941176	1,25387101145572	34					
	Andere	1,11538461538462	1,24344434783140	26					
	Gesamt	1,08333333333333	1,23907656521430	60					
Korretheit Typ K	Geographie-Hauptfach	1,20588235294118	1,12221133872816	34					
	Andere	1,07692307692308	1,38341828593024	26					
	Gesamt	1,15000000000000	1,23267656967419	60					

Tests der Innersubjekteffekte									
Maß: Korrektheit									
Quelle		Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat	Nichtzentralitäts-Parameter	Beobachtete Schärfe <sup>a</sup>
Visualisierungstyp	Sphärizität angenommen	.087	1	.087	.165	.686	.003	.165	.068
	Greenhouse-Geisser	.087	1.000	.087	.165	.686	.003	.165	.068
	Huynh-Feldt	.087	1.000	.087	.165	.686	.003	.165	.068
	Untergrenze	.087	1.000	.087	.165	.686	.003	.165	.068
Visualisierungstyp * Ausbildung	Sphärizität angenommen	.254	1	.254	.480	.491	.008	.480	.105
	Greenhouse-Geisser	.254	1.000	.254	.480	.491	.008	.480	.105
	Huynh-Feldt	.254	1.000	.254	.480	.491	.008	.480	.105
	Untergrenze	.254	1.000	.254	.480	.491	.008	.480	.105
Fehler(Visualisierungstyp)	Sphärizität angenommen	30.613	58	.528					
	Greenhouse-Geisser	30.613	58.000	.528					
	Huynh-Feldt	30.613	58.000	.528					
	Untergrenze	30.613	58.000	.528					

a. Unter Verwendung von Alpha = ,05 berechnet

### E5 <GLM repeated measures>: Ähnlichkeitswert bei Ähnlichkeitsaufgaben – Geographie-Studium

Deskriptive Statistiken					Innersubjektfaktoren				
					Maß: Ähnlichkeit				
					Visualisierungstyp Abhängige Variable				
					1 ÄhnlichkeitTypD				
					2 ÄhnlichkeitTypK				
					Zwischensubjektfaktoren				
					Wertelabel N				
					Ausbildung 1 Geographie-Hauptfach 34				
					2 Andere 26				
	Ausbildung	Mittelwert	Standardabweichung	N					
Ähnlichkeitswert Typ D	Geographie-Hauptfach	6,48529411764706	1,27483256117091	34					
	Andere	6,62179487179487	1,53388952296251	26					
	Gesamt	6,54444444444444	1,39168456860053	60					
Ähnlichkeitswert Typ K	Geographie-Hauptfach	6,28921568627451	1,24337220910965	34					
	Andere	6,41025641025641	1,64804987062539	26					
	Gesamt	6,34166666666667	1,42099703730784	60					

Tests der Innersubjekteffekte									
Maß: Ähnlichkeit									
Quelle		Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat	Nichtzentralitäts-Parameter	Beobachtete Schärfe <sup>a</sup>
Visualisierungstyp	Sphärizität angenommen	1.224	1	1.224	3.251	.077	.053	3.251	.426
	Greenhouse-Geisser	1.224	1.000	1.224	3.251	.077	.053	3.251	.426
	Huynh-Feldt	1.224	1.000	1.224	3.251	.077	.053	3.251	.426
	Untergrenze	1.224	1.000	1.224	3.251	.077	.053	3.251	.426
Visualisierungstyp * Ausbildung	Sphärizität angenommen	.002	1	.002	.005	.946	.000	.005	.051
	Greenhouse-Geisser	.002	1.000	.002	.005	.946	.000	.005	.051
	Huynh-Feldt	.002	1.000	.002	.005	.946	.000	.005	.051
	Untergrenze	.002	1.000	.002	.005	.946	.000	.005	.051
Fehler(Visualisierungstyp)	Sphärizität angenommen	21.834	58	.376					
	Greenhouse-Geisser	21.834	58.000	.376					
	Huynh-Feldt	21.834	58.000	.376					
	Untergrenze	21.834	58.000	.376					

a. Unter Verwendung von Alpha = ,05 berechnet

Tests der Zwischensubjekteffekte									
Maß: Ähnlichkeit									
Transformierte Variable: Mittel									
Quelle		Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat	Nichtzentralitäts-Parameter	Beobachtete Schärfe <sup>a</sup>
Konstanter Term		2453.021	1	2453.021	1348.065	.000	.959	1348.065	1.000
Ausbildung		.244	1	.244	.134	.715	.002	.134	.065
Fehler		105.540	58	1.820					

a. Unter Verwendung von Alpha = ,05 berechnet

## E6 <GLM repeated measures>: Einfluss verschiedener Faktoren bei Ähnlichkeitsaufgaben – Geographie-Studium

Deskriptive Statistiken				
	Ausbildung	Mittelwert	Standardabweichung	N
Regionzugehörigkeit Typ D	Geographie-Hauptfach	3,22549019607843	.508980549960185	34
	Andere	3,00000000000000	.839311887467712	26
	Gesamt	3,12777777777778	.675343719884608	60
Regionzugehörigkeit Typ K	Geographie-Hauptfach	3,14215686274510	.611522845147083	34
	Andere	3,00641025641026	.842589618353700	26
	Gesamt	3,08333333333333	.717352844546042	60
Regionsgrenzen Typ D	Geographie-Hauptfach	2,89215686274510	.699786346592688	34
	Andere	2,81410256410256	.901020314331362	26
	Gesamt	2,85833333333333	.787032975138247	60
Regionsgrenzen Typ K	Geographie-Hauptfach	2,91666666666667	.654240248708709	34
	Andere	2,78846153846154	.84532390889185	26
	Gesamt	2,86111111111111	.739118594202882	60
Begriffsdominanz Typ D	Geographie-Hauptfach	2,63235294117647	.766270661631230	34
	Andere	2,72435897435897	.847142002647574	26
	Gesamt	2,67222222222222	.796630537598159	60
Begriffsdominanz Typ K	Geographie-Hauptfach	2,62254901960784	.808528637792431	34
	Andere	2,75000000000000	1,00692049779955	26
	Gesamt	2,67777777777778	.894041094876679	60
Räumliche Distanz Typ D	Geographie-Hauptfach	2,42647058823529	.784091806096934	34
	Andere	2,04487179487180	.901304847069297	26
	Gesamt	2,26111111111111	.851146709979260	60
Räumliche Distanz Typ K	Geographie-Hauptfach	2,47549019607843	.896243895058577	34
	Andere	2,04487179487179	.935185890383744	26
	Gesamt	2,28888888888889	.930679574050321	60
Oberfläche Typ D	Geographie-Hauptfach	2,20588235294118	.942861558448467	34
	Andere	2,00641025641026	.917097800710785	26
	Gesamt	2,11944444444444	.929274703843919	60
Oberfläche Typ K	Geographie-Hauptfach	2,29411764705882	.977059099571998	34
	Andere	2,16025641025641	.998310538672277	26
	Gesamt	2,23611111111111	.980166626878895	60
Allgemein Typ D	Geographie-Hauptfach	2,67647058823529	.380548019841264	34
	Andere	2,51794871794872	.527739899675501	26
	Gesamt	2,60777777777778	.453085952221574	60
Allgemein Typ K	Geographie-Hauptfach	2,69019607843137	.396069018554824	34
	Andere	2,55000000000000	.514003891035956	26
	Gesamt	2,62944444444445	.452325719126596	60

Innersubjektfaktoren		
Maß	Visualisierungstyp	Abhängige Variable
Regionzugehörigkeit	1	RegionTypD
	2	RegionTypK
Regionsgrenzen	1	GrenzeTypD
	2	GrenzeTypK
Begriffsdominanz	1	DominanzTypD
	2	DominanzTypK
Distanz	1	DistanzTypD
	2	DistanzTypK
Oberflächenbeschaffenheit	1	OberflächeTypD
	2	OberflächeTypK
Allgemein	1	AllgemeinTypD
	2	AllgemeinTypK

Zwischensubjektfaktoren		
	Wertelabel	N
Ausbildung	1 Geographie-Hauptfach	34
	2 Andere	26

Tests der Zwischensubjekteffekte									
Transformierte Variable: Mittel									
Quelle	Maß	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat	Nichtzentralitäts-Parameter	Beobachtete Schärfe <sup>a</sup>
Konstanter Term	Regionzugehörigkeit	563,982	1	563,982	1291,038	.000	.957	1291,038	1,000
	Regionsgrenzen	479,643	1	479,643	865,892	.000	.937	865,892	1,000
	Begriffsdominanz	424,014	1	424,014	634,670	.000	.916	634,670	1,000
	Distanz	297,800	1	297,800	406,098	.000	.875	406,098	1,000
	Oberflächenbeschaffenheit	276,659	1	276,659	318,360	.000	.846	318,360	1,000
	Allgemein	401,046	1	401,046	2039,267	.000	.972	2039,267	1,000
Ausbildung	Regionzugehörigkeit	.481	1	.481	1,100	.299	.019	1,100	.178
	Regionsgrenzen	.157	1	.157	.283	.597	.005	.283	.082
	Begriffsdominanz	.177	1	.177	.266	.608	.005	.266	.080
	Distanz	2,430	1	2,430	3,314	.074	.054	3,314	.433
	Oberflächenbeschaffenheit	.409	1	.409	.471	.495	.008	.471	.104
	Allgemein	.329	1	.329	1,671	.201	.028	1,671	.246
Fehler	Regionzugehörigkeit	25,337	58	.437					
	Regionsgrenzen	32,128	58	.554					
	Begriffsdominanz	38,749	58	.668					
	Distanz	42,533	58	.733					
	Oberflächenbeschaffenheit	50,403	58	.869					
	Allgemein	11,406	58	.197					

a. Unter Verwendung von Alpha = ,05 berechnet

Tests auf Univariate

Quelle	Maß		Quadratsumme vom			F	Sig.	Partielles Eta-Quadrat	Nichtzentralitäts-Parameter	Beobachtete Schärfe <sup>a</sup>	
			Typ III	df	Mittel der Quadrate						
Visualisierungstyp	Regionzugehörigkeit	Sphärizität angenommen	.044	1	.044	.453	.503	.008	.453	.102	
		Greenhouse-Geisser	.044	1.000	.044	.453	.503	.008	.453	.102	
		Huynh-Feldt	.044	1.000	.044	.453	.503	.008	.453	.102	
		Untergrenze	.044	1.000	.044	.453	.503	.008	.453	.102	
	Regionsgrenzen	Sphärizität angenommen	9.427E-006	1	9.427E-006	.000	.991	.000	.000	.050	
		Greenhouse-Geisser	9.427E-006	1.000	9.427E-006	.000	.991	.000	.000	.050	
		Huynh-Feldt	9.427E-006	1.000	9.427E-006	.000	.991	.000	.000	.050	
		Untergrenze	9.427E-006	1.000	9.427E-006	.000	.991	.000	.000	.050	
	Begriffsdominanz	Sphärizität angenommen	.002	1	.002	.016	.900	.000	.016	.052	
		Greenhouse-Geisser	.002	1.000	.002	.016	.900	.000	.016	.052	
		Huynh-Feldt	.002	1.000	.002	.016	.900	.000	.016	.052	
		Untergrenze	.002	1.000	.002	.016	.900	.000	.016	.052	
	Distanz	Sphärizität angenommen	.018	1	.018	.263	.610	.005	.263	.080	
		Greenhouse-Geisser	.018	1.000	.018	.263	.610	.005	.263	.080	
		Huynh-Feldt	.018	1.000	.018	.263	.610	.005	.263	.080	
		Untergrenze	.018	1.000	.018	.263	.610	.005	.263	.080	
	Oberflächenbeschaffenheit	Sphärizität angenommen	.432	1	.432	4.190	.045	.067	4.190	.521	
		Greenhouse-Geisser	.432	1.000	.432	4.190	.045	.067	4.190	.521	
		Huynh-Feldt	.432	1.000	.432	4.190	.045	.067	4.190	.521	
		Untergrenze	.432	1.000	.432	4.190	.045	.067	4.190	.521	
	Allgemein	Sphärizität angenommen	.015	1	.015	1.260	.266	.021	1.260	.197	
		Greenhouse-Geisser	.015	1.000	.015	1.260	.266	.021	1.260	.197	
		Huynh-Feldt	.015	1.000	.015	1.260	.266	.021	1.260	.197	
		Untergrenze	.015	1.000	.015	1.260	.266	.021	1.260	.197	
	Visualisierungstyp * Ausbildung	Regionzugehörigkeit	Sphärizität angenommen	.059	1	.059	.617	.435	.011	.617	.121
			Greenhouse-Geisser	.059	1.000	.059	.617	.435	.011	.617	.121
			Huynh-Feldt	.059	1.000	.059	.617	.435	.011	.617	.121
			Untergrenze	.059	1.000	.059	.617	.435	.011	.617	.121
Regionsgrenzen		Sphärizität angenommen	.019	1	.019	.256	.614	.004	.256	.079	
		Greenhouse-Geisser	.019	1.000	.019	.256	.614	.004	.256	.079	
		Huynh-Feldt	.019	1.000	.019	.256	.614	.004	.256	.079	
		Untergrenze	.019	1.000	.019	.256	.614	.004	.256	.079	
Begriffsdominanz		Sphärizität angenommen	.009	1	.009	.080	.779	.001	.080	.059	
		Greenhouse-Geisser	.009	1.000	.009	.080	.779	.001	.080	.059	
		Huynh-Feldt	.009	1.000	.009	.080	.779	.001	.080	.059	
		Untergrenze	.009	1.000	.009	.080	.779	.001	.080	.059	
Distanz		Sphärizität angenommen	.018	1	.018	.263	.610	.005	.263	.080	
		Greenhouse-Geisser	.018	1.000	.018	.263	.610	.005	.263	.080	
		Huynh-Feldt	.018	1.000	.018	.263	.610	.005	.263	.080	
		Untergrenze	.018	1.000	.018	.263	.610	.005	.263	.080	
Oberflächenbeschaffenheit		Sphärizität angenommen	.032	1	.032	.308	.581	.005	.308	.085	
		Greenhouse-Geisser	.032	1.000	.032	.308	.581	.005	.308	.085	
		Huynh-Feldt	.032	1.000	.032	.308	.581	.005	.308	.085	
		Untergrenze	.032	1.000	.032	.308	.581	.005	.308	.085	
Allgemein		Sphärizität angenommen	.002	1	.002	.202	.655	.003	.202	.073	
		Greenhouse-Geisser	.002	1.000	.002	.202	.655	.003	.202	.073	
		Huynh-Feldt	.002	1.000	.002	.202	.655	.003	.202	.073	
		Untergrenze	.002	1.000	.002	.202	.655	.003	.202	.073	
Fehler(Visualisierungstyp)		Regionzugehörigkeit	Sphärizität angenommen	5.576	58	.096					
			Greenhouse-Geisser	5.576	58.000	.096					
			Huynh-Feldt	5.576	58.000	.096					
			Untergrenze	5.576	58.000	.096					
	Regionsgrenzen	Sphärizität angenommen	4.190	58	.072						
		Greenhouse-Geisser	4.190	58.000	.072						
		Huynh-Feldt	4.190	58.000	.072						
		Untergrenze	4.190	58.000	.072						
	Begriffsdominanz	Sphärizität angenommen	6.740	58	.116						
		Greenhouse-Geisser	6.740	58.000	.116						
		Huynh-Feldt	6.740	58.000	.116						
		Untergrenze	6.740	58.000	.116						
	Distanz	Sphärizität angenommen	3.904	58	.067						
		Greenhouse-Geisser	3.904	58.000	.067						
		Huynh-Feldt	3.904	58.000	.067						
		Untergrenze	3.904	58.000	.067						
	Oberflächenbeschaffenheit	Sphärizität angenommen	5.977	58	.103						
		Greenhouse-Geisser	5.977	58.000	.103						
		Huynh-Feldt	5.977	58.000	.103						
		Untergrenze	5.977	58.000	.103						
	Allgemein	Sphärizität angenommen	.711	58	.012						
		Greenhouse-Geisser	.711	58.000	.012						
		Huynh-Feldt	.711	58.000	.012						
		Untergrenze	.711	58.000	.012						

a. Unter Verwendung von Alpha = ,05 berechnet

E7 <GLM repeated measures>: Sicherheit bei Ähnlichkeitsaufgaben – Geographic-Studium

Innersubjektffaktoren

Maß: Sicherheit	
Visualisierungstyp	Abhängige Variable
1	ÄhnSicher TypD
2	ÄhnSicher TypK

Zwischensubjektffaktoren

		Wertelabel	N
Ausbildung	1	Geographie-Hauptfach	34
	2	Andere	26

Deskriptive Statistiken

		Ausbildung	Mittelwert	Standardabweichung	N
Ähnlichkeit Sicherheit Typ D	Geographie-Hauptfach	3,22549019607843	.758656389109154	34	
	Andere	3,60256410256410	1,02640356269075	26	
	Gesamt	3,38888888888889	.896565253052991	60	
Ähnlichkeit Sicherheit Typ K	Geographie-Hauptfach	3,11274509803922	.818268464415904	34	
	Andere	3,60256410256410	1,01003511828596	26	
	Gesamt	3,32500000000000	.930961979436591	60	

Bei dieser Auswertung ist zu beachten, dass ursprünglich der Wert 1 zur Einschätzung «Sehr sicher» zugeordnet wurde. Dies wurde für die Diagramme im Kapitel Resultate umgewandelt. Doch hier im Anhang gilt: je tiefer der Sicherheitswert, desto sicherer fühlten sich die Testpersonen.

Tests der Innersubjektffekte

Maß: Sicherheit		Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat	Nichtzentralitäts-Parameter	Beobachtete Schärfe <sup>a</sup>
Visualisierungstyp	Sphärizität angenommen	.094	1	.094	1.159	.286	.020	1.159	.185
	Greenhouse-Geisser	.094	1.000	.094	1.159	.286	.020	1.159	.185
	Huynh-Feldt	.094	1.000	.094	1.159	.286	.020	1.159	.185
	Untergrenze	.094	1.000	.094	1.159	.286	.020	1.159	.185
Visualisierungstyp * Ausbildung	Sphärizität angenommen	.094	1	.094	1.159	.286	.020	1.159	.185
	Greenhouse-Geisser	.094	1.000	.094	1.159	.286	.020	1.159	.185
	Huynh-Feldt	.094	1.000	.094	1.159	.286	.020	1.159	.185
	Untergrenze	.094	1.000	.094	1.159	.286	.020	1.159	.185
Fehler(Visualisierungstyp)	Sphärizität angenommen	4.687	58	.081					
	Greenhouse-Geisser	4.687	58.000	.081					
	Huynh-Feldt	4.687	58.000	.081					
	Untergrenze	4.687	58.000	.081					

a. Unter Verwendung von Alpha = ,05 berechnet

Tests der Zwischensubjektffekte

Maß: Sicherheit		Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat	Nichtzentralitäts-Parameter	Beobachtete Schärfe <sup>a</sup>
Transformierte Variable: Mittel									
Konstanter Term	675.607	1	675.607	888.108	.000	.939	888.108	1.000	
Ausbildung	2.768	1	2.768	3.639	.061	.059	3.639	.467	
Fehler	44.122	58	.761						

a. Unter Verwendung von Alpha = ,05 berechnet

## E8 <GLM repeated measures>: Unterschiede zwischen Ähnlichkeitsaufgabentypen – Geographie-Studium

Deskriptive Statistiken				
	Ausbildung	Mittelwert	Standardabweichung	N
Aufgabe1 Typ D	Geographie-Hauptfach	6.353	2.1445	34
	Andere	6.692	2.3962	26
	Gesamt	6.500	2.2436	60
Aufgabe2 Typ D	Geographie-Hauptfach	6.56	2.077	34
	Andere	7.77	1.557	26
	Gesamt	7.08	1.951	60
Aufgabe3 Typ D	Geographie-Hauptfach	6.62	2.075	34
	Andere	6.88	2.304	26
	Gesamt	6.73	2.162	60
Aufgabe4 Typ D	Geographie-Hauptfach	5.41	2.148	34
	Andere	5.54	2.832	26
	Gesamt	5.47	2.446	60
Aufgabe5 Typ D	Geographie-Hauptfach	6.76	1.577	34
	Andere	6.50	1.772	26
	Gesamt	6.65	1.655	60
Aufgabe6 Typ D	Geographie-Hauptfach	7.21	1.409	34
	Andere	6.35	2.607	26
	Gesamt	6.83	2.043	60
Aufgabe1 Typ K	Geographie-Hauptfach	5.88	2.253	34
	Andere	6.38	2.099	26
	Gesamt	6.10	2.184	60
Aufgabe2 Typ K	Geographie-Hauptfach	6.79	2.320	34
	Andere	7.54	1.655	26
	Gesamt	7.12	2.076	60
Aufgabe3 Typ K	Geographie-Hauptfach	6.94	1.825	34
	Andere	6.54	2.370	26
	Gesamt	6.77	2.070	60
Aufgabe4 Typ K	Geographie-Hauptfach	5.24	2.230	34
	Andere	5.00	3.124	26
	Gesamt	5.13	2.633	60
Aufgabe5 Typ K	Geographie-Hauptfach	6.32	1.736	34
	Andere	6.15	1.891	26
	Gesamt	6.25	1.791	60
Aufgabe6 Typ K	Geographie-Hauptfach	6.56	1.862	34
	Andere	6.85	2.130	26
	Gesamt	6.68	1.970	60

Innersubjektfaktoren		
Maß: Ähnlichkeit		
Visualisierung	Aufgabe	Abhängige Variable
1	1	Aufg1TypD
	2	Aufg2TypD
	3	Aufg3TypD
	4	Aufg4TypD
	5	Aufg5TypD
	6	Aufg6TypD
2	1	Aufg1TypK
	2	Aufg2TypK
	3	Aufg3TypK
	4	Aufg4TypK
	5	Aufg5TypK
	6	Aufg6TypK

Zwischensubjektfaktoren			
Wertelabel			
Ausbildung	1.0	Geographie-Hauptfach	N
	2.0	Andere	26

Tests der Innersubjektffekte

Maß: Ähnlichkeit										
Quelle	Quadratsumme vom Typ III			df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat	Nichtzentralitäts-Parameter	Beobachtete Schärfe <sup>a</sup>
Visualisierung	Sphärität angenommen	7.344	1	7.344	3.251	.077	.053	3.251	.426	
	Greenhouse-Geisser	7.344	1.000	7.344	3.251	.077	.053	3.251	.426	
	Huynh-Feldt	7.344	1.000	7.344	3.251	.077	.053	3.251	.426	
	Untergrenze	7.344	1.000	7.344	3.251	.077	.053	3.251	.426	
Visualisierung * Ausbildung	Sphärität angenommen	.011	1	.011	.005	.946	.000	.005	.051	
	Greenhouse-Geisser	.011	1.000	.011	.005	.946	.000	.005	.051	
	Huynh-Feldt	.011	1.000	.011	.005	.946	.000	.005	.051	
	Untergrenze	.011	1.000	.011	.005	.946	.000	.005	.051	
Fehler(Visualisierung)	Sphärität angenommen	131.005	58	2.259						
	Greenhouse-Geisser	131.005	58.000	2.259						
	Huynh-Feldt	131.005	58.000	2.259						
	Untergrenze	131.005	58.000	2.259						
Aufgabe	Sphärität angenommen	239.072	5	47.814	11.933	.000	.171	59.664	1.000	
	Greenhouse-Geisser	239.072	3.493	68.444	11.933	.000	.171	41.681	1.000	
	Huynh-Feldt	239.072	3.808	62.777	11.933	.000	.171	45.444	1.000	
	Untergrenze	239.072	1.000	239.072	11.933	.001	.171	11.933	.925	
Aufgabe * Ausbildung	Sphärität angenommen	34.461	5	6.892	1.720	.130	.029	8.600	.591	
	Greenhouse-Geisser	34.461	3.493	9.866	1.720	.155	.029	6.008	.484	
	Huynh-Feldt	34.461	3.808	9.049	1.720	.150	.029	6.551	.508	
	Untergrenze	34.461	1.000	34.461	1.720	.195	.029	1.720	.252	
Fehler(Aufgabe)	Sphärität angenommen	1162.015	290	4.007						
	Greenhouse-Geisser	1162.015	202.593	5.736						
	Huynh-Feldt	1162.015	220.881	5.261						
	Untergrenze	1162.015	58.000	20.035						
Visualisierung * Aufgabe	Sphärität angenommen	5.613	5	1.123	.580	.715	.010	2.902	.212	
	Greenhouse-Geisser	5.613	3.912	1.435	.580	.673	.010	2.270	.189	
	Huynh-Feldt	5.613	4.302	1.305	.580	.690	.010	2.497	.197	
	Untergrenze	5.613	1.000	5.613	.580	.449	.010	.580	.116	
Visualisierung * Aufgabe * Ausbildung	Sphärität angenommen	15.813	5	3.163	1.635	.151	.027	8.174	.566	
	Greenhouse-Geisser	15.813	3.912	4.042	1.635	.168	.027	6.395	.493	
	Huynh-Feldt	15.813	4.302	3.675	1.635	.161	.027	7.034	.520	
	Untergrenze	15.813	1.000	15.813	1.635	.206	.027	1.635	.242	
Fehler(Visualisierung* Aufgabe)	Sphärität angenommen	560.996	290	1.934						
	Greenhouse-Geisser	560.996	226.887	2.473						
	Huynh-Feldt	560.996	249.541	2.248						
	Untergrenze	560.996	58.000	9.672						

a. Unter Verwendung von Alpha = .05 berechnet

Tests der Zwischensubjektffekte

Maß: Ähnlichkeit										
Transformierte Variable: Mittel										
Quelle	Quadratsumme vom Typ III			df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat	Nichtzentralitäts-Parameter	Beobachtete Schärfe <sup>a</sup>
Konstanter Term	2453.021	1	2453.021	1348.065	.000		.959	1348.065	1.000	
Ausbildung	.244	1	.244	.194	.715		.002	.194	.065	
Fehler	105.540	58	1.820							

a. Unter Verwendung von Alpha = .05 berechnet

## E9 <GLM repeated measures>: Unterschiede in Ähnlichkeitsaufgabentypen – Geographie-Studium

Deskriptive Statistiken				
	Ausbildung	Mittelwert	Standardabweichung	N
Aufgabe1 Typ D	Geographie-Hauptfach	6.353	2.1445	34
	Andere	6.692	2.3962	26
	Gesamt	6.500	2.2436	60
Aufgabe1 Typ K	Geographie-Hauptfach	5.88	2.253	34
	Andere	6.38	2.099	26
	Gesamt	6.10	2.184	60
Aufgabe2 Typ D	Geographie-Hauptfach	6.56	2.077	34
	Andere	7.77	1.557	26
	Gesamt	7.08	1.951	60
Aufgabe2 Typ K	Geographie-Hauptfach	6.79	2.320	34
	Andere	7.54	1.655	26
	Gesamt	7.12	2.076	60
Aufgabe3 Typ D	Geographie-Hauptfach	6.62	2.075	34
	Andere	6.88	2.304	26
	Gesamt	6.73	2.162	60
Aufgabe3 Typ K	Geographie-Hauptfach	6.94	1.825	34
	Andere	6.54	2.370	26
	Gesamt	6.77	2.070	60
Aufgabe4 Typ D	Geographie-Hauptfach	5.41	2.148	34
	Andere	5.54	2.832	26
	Gesamt	5.47	2.446	60
Aufgabe4 Typ K	Geographie-Hauptfach	5.24	2.230	34
	Andere	5.00	3.124	26
	Gesamt	5.13	2.633	60
Aufgabe5 Typ D	Geographie-Hauptfach	6.76	1.577	34
	Andere	6.50	1.772	26
	Gesamt	6.65	1.655	60
Aufgabe5 Typ K	Geographie-Hauptfach	6.32	1.736	34
	Andere	6.15	1.891	26
	Gesamt	6.25	1.791	60
Aufgabe6 Typ D	Geographie-Hauptfach	7.21	1.409	34
	Andere	6.35	2.607	26
	Gesamt	6.83	2.043	60
Aufgabe6 Typ K	Geographie-Hauptfach	6.56	1.862	34
	Andere	6.85	2.130	26
	Gesamt	6.68	1.970	60

Innersubjektfaktoren		
Maß	Visualisierung	Abhängige Variable
Aufgabe1	1	Aufg1TypD
	2	Aufg1TypK
Aufgabe2	1	Aufg2TypD
	2	Aufg2TypK
Aufgabe3	1	Aufg3TypD
	2	Aufg3TypK
Aufgabe4	1	Aufg4TypD
	2	Aufg4TypK
Aufgabe5	1	Aufg5TypD
	2	Aufg5TypK
Aufgabe6	1	Aufg6TypD
	2	Aufg6TypK

Zwischensubjektfaktoren			
	Wertelabel		N
Ausbildung	1.0	Geographie-Hauptfach	34
	2.0	Andere	26

Tests der Zwischensubjekteffekte

Transformierte Variable: Mittel										
Quelle	Maß	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat	Nichtzentralitäts-Parameter	Beobachtete Schärfe <sup>a</sup>	
Konstanter Term	Aufgabe1	2359.942	1	2359.942	612.448	.000	.913	612.448	1.000	
	Aufgabe2	3025.608	1	3025.608	962.482	.000	.943	962.482	1.000	
	Aufgabe3	2681.551	1	2681.551	809.284	.000	.933	809.284	1.000	
	Aufgabe4	1653.177	1	1653.177	325.521	.000	.849	325.521	1.000	
	Aufgabe5	2440.778	1	2440.778	927.354	.000	.941	927.354	1.000	
	Aufgabe6	2676.607	1	2676.607	918.387	.000	.941	918.387	1.000	
Ausbildung	Aufgabe1	2.609	1	2.609	.677	.414	.012	.677	.128	
	Aufgabe2	14.074	1	14.074	4.477	.039	.072	4.477	.548	
	Aufgabe3	.068	1	.068	.020	.887	.000	.020	.052	
	Aufgabe4	.043	1	.043	.009	.927	.000	.009	.051	
	Aufgabe5	.695	1	.695	.264	.609	.005	.264	.080	
	Aufgabe6	1.207	1	1.207	.414	.522	.007	.414	.097	
Fehler	Aufgabe1	223.491	58	3.853						
	Aufgabe2	182.326	58	3.144						
	Aufgabe3	192.182	58	3.313						
	Aufgabe4	294.557	58	5.079						
	Aufgabe5	152.655	58	2.632						
	Aufgabe6	169.039	58	2.914						

a. Unter Verwendung von Alpha = ,05 berechnet

Tests auf Univariate

Quelle	Wk6		Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat	Nichtzentralitäts-Parameter	Beobachtete Schärfe <sup>a</sup>	
Visualisierung	Aufgabe1	Sphärizität angenommen	4.462	1	4.462	2.054	.157	.034	2.054	.291	
		Greenhouse-Geisser	4.462	1.000	4.462	2.054	.157	.034	2.054	.291	
		Huynh-Feldt	4.462	1.000	4.462	2.054	.157	.034	2.054	.291	
		Untergrenze	4.462	1.000	4.462	2.054	.157	.034	2.054	.291	
	Aufgabe2	Sphärizität angenommen	.000	1	.000	.000	.992	.000	.000	.050	
		Greenhouse-Geisser	.000	1.000	.000	.000	.992	.000	.000	.050	
		Huynh-Feldt	.000	1.000	.000	.000	.992	.000	.000	.050	
		Untergrenze	.000	1.000	.000	.000	.992	.000	.000	.050	
	Aufgabe3	Sphärizität angenommen	.004	1	.004	.002	.969	.000	.002	.050	
		Greenhouse-Geisser	.004	1.000	.004	.002	.969	.000	.002	.050	
		Huynh-Feldt	.004	1.000	.004	.002	.969	.000	.002	.050	
		Untergrenze	.004	1.000	.004	.002	.969	.000	.002	.050	
	Aufgabe4	Sphärizität angenommen	3.765	1	3.765	1.272	.264	.021	1.272	.199	
		Greenhouse-Geisser	3.765	1.000	3.765	1.272	.264	.021	1.272	.199	
		Huynh-Feldt	3.765	1.000	3.765	1.272	.264	.021	1.272	.199	
		Untergrenze	3.765	1.000	3.765	1.272	.264	.021	1.272	.199	
	Aufgabe5	Sphärizität angenommen	4.567	1	4.567	6.001	.017	.094	6.001	.673	
		Greenhouse-Geisser	4.567	1.000	4.567	6.001	.017	.094	6.001	.673	
		Huynh-Feldt	4.567	1.000	4.567	6.001	.017	.094	6.001	.673	
		Untergrenze	4.567	1.000	4.567	6.001	.017	.094	6.001	.673	
	Aufgabe6	Sphärizität angenommen	.159	1	.159	.074	.787	.001	.074	.058	
		Greenhouse-Geisser	.159	1.000	.159	.074	.787	.001	.074	.058	
		Huynh-Feldt	.159	1.000	.159	.074	.787	.001	.074	.058	
		Untergrenze	.159	1.000	.159	.074	.787	.001	.074	.058	
	Visualisierung * Ausbildung	Aufgabe1	Sphärizität angenommen	.195	1	.195	.090	.765	.002	.090	.060
			Greenhouse-Geisser	.195	1.000	.195	.090	.765	.002	.090	.060
			Huynh-Feldt	.195	1.000	.195	.090	.765	.002	.090	.060
			Untergrenze	.195	1.000	.195	.090	.765	.002	.090	.060
Aufgabe2		Sphärizität angenommen	1.600	1	1.600	1.100	.299	.019	1.100	.178	
		Greenhouse-Geisser	1.600	1.000	1.600	1.100	.299	.019	1.100	.178	
		Huynh-Feldt	1.600	1.000	1.600	1.100	.299	.019	1.100	.178	
		Untergrenze	1.600	1.000	1.600	1.100	.299	.019	1.100	.178	
Aufgabe3		Sphärizität angenommen	3.304	1	3.304	1.362	.248	.023	1.362	.209	
		Greenhouse-Geisser	3.304	1.000	3.304	1.362	.248	.023	1.362	.209	
		Huynh-Feldt	3.304	1.000	3.304	1.362	.248	.023	1.362	.209	
		Untergrenze	3.304	1.000	3.304	1.362	.248	.023	1.362	.209	
Aufgabe4		Sphärizität angenommen	.965	1	.965	.326	.570	.006	.326	.087	
		Greenhouse-Geisser	.965	1.000	.965	.326	.570	.006	.326	.087	
		Huynh-Feldt	.965	1.000	.965	.326	.570	.006	.326	.087	
		Untergrenze	.965	1.000	.965	.326	.570	.006	.326	.087	
Aufgabe5		Sphärizität angenommen	.067	1	.067	.087	.769	.002	.087	.060	
		Greenhouse-Geisser	.067	1.000	.067	.087	.769	.002	.087	.060	
		Huynh-Feldt	.067	1.000	.067	.087	.769	.002	.087	.060	
		Untergrenze	.067	1.000	.067	.087	.769	.002	.087	.060	
Aufgabe6		Sphärizität angenommen	9.693	1	9.693	4.493	.038	.072	4.493	.550	
		Greenhouse-Geisser	9.693	1.000	9.693	4.493	.038	.072	4.493	.550	
		Huynh-Feldt	9.693	1.000	9.693	4.493	.038	.072	4.493	.550	
		Untergrenze	9.693	1.000	9.693	4.493	.038	.072	4.493	.550	
Fehler(Visualisierung)		Aufgabe1	Sphärizität angenommen	126.005	58	2.172					
			Greenhouse-Geisser	126.005	58.000	2.172					
			Huynh-Feldt	126.005	58.000	2.172					
			Untergrenze	126.005	58.000	2.172					
	Aufgabe2	Sphärizität angenommen	84.367	58	1.455						
		Greenhouse-Geisser	84.367	58.000	1.455						
		Huynh-Feldt	84.367	58.000	1.455						
		Untergrenze	84.367	58.000	1.455						
	Aufgabe3	Sphärizität angenommen	140.663	58	2.425						
		Greenhouse-Geisser	140.663	58.000	2.425						
		Huynh-Feldt	140.663	58.000	2.425						
		Untergrenze	140.663	58.000	2.425						
	Aufgabe4	Sphärizität angenommen	171.701	58	2.960						
		Greenhouse-Geisser	171.701	58.000	2.960						
		Huynh-Feldt	171.701	58.000	2.960						
		Untergrenze	171.701	58.000	2.960						
	Aufgabe5	Sphärizität angenommen	44.133	58	.761						
		Greenhouse-Geisser	44.133	58.000	.761						
		Huynh-Feldt	44.133	58.000	.761						
		Untergrenze	44.133	58.000	.761						
	Aufgabe6	Sphärizität angenommen	125.132	58	2.157						
		Greenhouse-Geisser	125.132	58.000	2.157						
		Huynh-Feldt	125.132	58.000	2.157						
		Untergrenze	125.132	58.000	2.157						

a. Unter Verwendung von Alpha = .05 berechnet

## E10 <GLM repeated measures>: Einfluss verschiedener Faktoren bei Ähnlichkeitsaufgabe 5 – Geographie-Studium

Deskriptive Statistiken				
	Ausbildung	Mittelwert	Standardabweichung	N
Aufg.5 Regionszugehörigkeit Typ D	Geographie-Hauptfach	2.529	1.1867	34
	Andere	2.692	1.3197	26
	<b>Gesamt</b>	<b>2.600</b>	<b>1.2378</b>	<b>60</b>
Aufg.5 Regionszugehörigkeit Typ K	Geographie-Hauptfach	2.324	1.2240	34
	Andere	2.538	1.2722	26
	<b>Gesamt</b>	<b>2.417</b>	<b>1.2391</b>	<b>60</b>
Aufg.5 Regionsgrenzen Typ D	Geographie-Hauptfach	2.059	1.1791	34
	Andere	2.154	1.1556	26
	<b>Gesamt</b>	<b>2.100</b>	<b>1.1601</b>	<b>60</b>
Aufg.5 Regionsgrenzen Typ K	Geographie-Hauptfach	1.853	1.0483	34
	Andere	2.192	1.2967	26
	<b>Gesamt</b>	<b>2.000</b>	<b>1.1644</b>	<b>60</b>
Aufg.5 Begriffsdominanz Typ D	Geographie-Hauptfach	3.206	1.0668	34
	Andere	3.269	1.0792	26
	<b>Gesamt</b>	<b>3.233</b>	<b>1.0635</b>	<b>60</b>
Aufg.5 Begriffsdominanz Typ K	Geographie-Hauptfach	3.206	1.1222	34
	Andere	3.423	1.0266	26
	<b>Gesamt</b>	<b>3.300</b>	<b>1.0783</b>	<b>60</b>
Aufg.5 Distanz Typ D	Geographie-Hauptfach	2.500	1.0225	34
	Andere	1.885	1.0706	26
	<b>Gesamt</b>	<b>2.233</b>	<b>1.0793</b>	<b>60</b>
Aufg.5 Distanz Typ K	Geographie-Hauptfach	2.500	1.0801	34
	Andere	1.885	1.1429	26
	<b>Gesamt</b>	<b>2.233</b>	<b>1.1404</b>	<b>60</b>
Aufg.5 Oberflächenbeschaffenheit Typ D	Geographie-Hauptfach	2.559	1.2108	34
	Andere	2.269	1.2824	26
	<b>Gesamt</b>	<b>2.433</b>	<b>1.2401</b>	<b>60</b>
Aufg.5 Oberflächenbeschaffenheit Typ K	Geographie-Hauptfach	2.529	1.2610	34
	Andere	2.231	1.3945	26
	<b>Gesamt</b>	<b>2.400</b>	<b>1.3174</b>	<b>60</b>

Innersubjektfaktoren		
Maß	Visualisierung	Abhängige Variable
Regionszugehörigkeit	1	Auf5RegionTypD
	2	Auf5RegionTypK
Regionsgrenzen	1	Auf5GrenzeTypD
	2	Auf5GrenzeTypK
Begriffsdominanz	1	Auf5DominanzTypD
	2	Auf5DominanzTypK
Distanz	1	Auf5DistanzTypD
	2	Auf5DistanzTypK
Oberflächenbeschaffenheit	1	Auf5OberflächeTypD
	2	Auf5OberflächeTypK

Zwischensubjektfaktoren		
	Wertelabel	N
Ausbildung	1.00	34
	2.00	26

Tests der Zwischensubjekteffekte

Transformierte Variable: Mittel										
Quelle	Maß	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat	Nichtzentralitäts-Parameter	Beobachtete Schärfe <sup>a</sup>	
Konstanter Term	Regionszugehörigkeit	749.052	1	749.052	313.818	.000	.844	313.818	1.000	
	Regionsgrenzen	502.357	1	502.357	204.741	.000	.779	204.741	1.000	
	Begriffsdominanz	1264.980	1	1264.980	661.655	.000	.919	661.655	1.000	
	Distanz	566.492	1	566.492	270.853	.000	.824	270.853	1.000	
	Oberflächenbeschaffenheit	677.249	1	677.249	231.582	.000	.800	231.582	1.000	
Ausbildung	Regionszugehörigkeit	1.052	1	1.052	.441	.509	.008	.441	.100	
	Regionsgrenzen	1.390	1	1.390	.567	.455	.010	.567	.115	
	Begriffsdominanz	.580	1	.580	.303	.584	.005	.303	.084	
	Distanz	11.159	1	11.159	5.335	.024	.084	5.335	.622	
	Oberflächenbeschaffenheit	2.549	1	2.549	.872	.354	.015	.872	.151	
Fehler	Regionszugehörigkeit	138.440	58	2.387						
	Regionsgrenzen	142.310	58	2.454						
	Begriffsdominanz	110.887	58	1.912						
	Distanz	121.308	58	2.092						
	Oberflächenbeschaffenheit	169.618	58	2.924						

a. Unter Verwendung von Alpha = .05 berechnet

Tests auf Univariate

Quelle	Maß		Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles B-Quadrat	Nichtzentralitäts-Parameter	Beobachtete Schärfe <sup>a</sup>
Visualisierung	Regionszugehörigkeit	Sphärizität angenommen	.953	1	.953	1.333	.253	.022	1.333	.206
		Greenhouse-Geisser	.953	1.000	.953	1.333	.253	.022	1.333	.206
		Huynh-Feldt	.953	1.000	.953	1.333	.253	.022	1.333	.206
		Untergrenze	.953	1.000	.953	1.333	.253	.022	1.333	.206
	Regionsgrenzen	Sphärizität angenommen	.206	1	.206	.785	.379	.013	.785	.140
		Greenhouse-Geisser	.206	1.000	.206	.785	.379	.013	.785	.140
		Huynh-Feldt	.206	1.000	.206	.785	.379	.013	.785	.140
		Untergrenze	.206	1.000	.206	.785	.379	.013	.785	.140
	Begriffsdominanz	Sphärizität angenommen	.174	1	.174	.427	.516	.007	.427	.098
		Greenhouse-Geisser	.174	1.000	.174	.427	.516	.007	.427	.098
		Huynh-Feldt	.174	1.000	.174	.427	.516	.007	.427	.098
		Untergrenze	.174	1.000	.174	.427	.516	.007	.427	.098
	Distanz	Sphärizität angenommen	.000	1	.000	.000	1.000	.000	.000	.050
		Greenhouse-Geisser	.000	1.000	.000	.000	1.000	.000	.000	.050
		Huynh-Feldt	.000	1.000	.000	.000	1.000	.000	.000	.050
		Untergrenze	.000	1.000	.000	.000	1.000	.000	.000	.050
	Oberflächenbeschaffenheit	Sphärizität angenommen	.034	1	.034	.094	.760	.002	.094	.060
		Greenhouse-Geisser	.034	1.000	.034	.094	.760	.002	.094	.060
		Huynh-Feldt	.034	1.000	.034	.094	.760	.002	.094	.060
		Untergrenze	.034	1.000	.034	.094	.760	.002	.094	.060
Visualisierung * Ausbildung	Regionszugehörigkeit	Sphärizität angenommen	.020	1	.020	.028	.868	.000	.028	.053
		Greenhouse-Geisser	.020	1.000	.020	.028	.868	.000	.028	.053
		Huynh-Feldt	.020	1.000	.020	.028	.868	.000	.028	.053
		Untergrenze	.020	1.000	.020	.028	.868	.000	.028	.053
	Regionsgrenzen	Sphärizität angenommen	.440	1	.440	1.672	.201	.028	1.672	.246
		Greenhouse-Geisser	.440	1.000	.440	1.672	.201	.028	1.672	.246
		Huynh-Feldt	.440	1.000	.440	1.672	.201	.028	1.672	.246
		Untergrenze	.440	1.000	.440	1.672	.201	.028	1.672	.246
	Begriffsdominanz	Sphärizität angenommen	.174	1	.174	.427	.516	.007	.427	.098
		Greenhouse-Geisser	.174	1.000	.174	.427	.516	.007	.427	.098
		Huynh-Feldt	.174	1.000	.174	.427	.516	.007	.427	.098
		Untergrenze	.174	1.000	.174	.427	.516	.007	.427	.098
	Distanz	Sphärizität angenommen	.000	1	.000	.000	1.000	.000	.000	.050
		Greenhouse-Geisser	.000	1.000	.000	.000	1.000	.000	.000	.050
		Huynh-Feldt	.000	1.000	.000	.000	1.000	.000	.000	.050
		Untergrenze	.000	1.000	.000	.000	1.000	.000	.000	.050
	Oberflächenbeschaffenheit	Sphärizität angenommen	.001	1	.001	.002	.968	.000	.002	.050
		Greenhouse-Geisser	.001	1.000	.001	.002	.968	.000	.002	.050
		Huynh-Feldt	.001	1.000	.001	.002	.968	.000	.002	.050
		Untergrenze	.001	1.000	.001	.002	.968	.000	.002	.050
Fehler(Visualisierung)	Regionszugehörigkeit	Sphärizität angenommen	41.472	58	.715					
		Greenhouse-Geisser	41.472	58.000	.715					
		Huynh-Feldt	41.472	58.000	.715					
		Untergrenze	41.472	58.000	.715					
	Regionsgrenzen	Sphärizität angenommen	15.260	58	.263					
		Greenhouse-Geisser	15.260	58.000	.263					
		Huynh-Feldt	15.260	58.000	.263					
		Untergrenze	15.260	58.000	.263					
	Begriffsdominanz	Sphärizität angenommen	23.692	58	.408					
		Greenhouse-Geisser	23.692	58.000	.408					
		Huynh-Feldt	23.692	58.000	.408					
		Untergrenze	23.692	58.000	.408					
	Distanz	Sphärizität angenommen	13.000	58	.224					
		Greenhouse-Geisser	13.000	58.000	.224					
		Huynh-Feldt	13.000	58.000	.224					
		Untergrenze	13.000	58.000	.224					
	Oberflächenbeschaffenheit	Sphärizität angenommen	20.966	58	.361					
		Greenhouse-Geisser	20.966	58.000	.361					
		Huynh-Feldt	20.966	58.000	.361					
		Untergrenze	20.966	58.000	.361					

a. Unter Verwendung von Alpha = .05 berechnet

E11 <GLM repeated measures>: Unterschiede in der Präferenz – Geographic-Studium

Deskriptive Statistiken					Innersubjektfaktoren	
Präferenz	Ausbildung	Mittelwert	Standardabweichun		Maß: Präferenz	
			g	N	Aufgabentyp	Abhängige Variable
Präferenz Probeaufgaben	Geographie-Hauptfach	3.848	1.4169	33	1	PräferenzProbe
	Andere	2.545	1.6250	22	2	PräferenzÜbersicht
	Gesamt	3.327	1.6222	55	3	PräferenzÄhnlichkeit
Präferenz Übersicht	Geographie-Hauptfach	2.424	1.1734	33	4	PräferenzInsgesamt
	Andere	2.545	1.4385	22		
	Gesamt	2.473	1.2745	55		
Präferenz Ähnlichkeit	Geographie-Hauptfach	2.909	1.1555	33	Zwischensubjektfaktoren	
	Andere	2.318	1.0414	22		
	Gesamt	2.673	1.1394	55		
Präferenz Insgesamt	Geographie-Hauptfach	3.212	1.3407	33	Wertelabel	
	Andere	1.773	.9726	22	Ausbildung 1	Geographie-Hauptfach
	Gesamt	2.636	1.3926	55	2	Andere
					N	
						33
						22

Tests der Innersubjekteffekte									
Maß: Präferenz									
Quelle		Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat	Nichtzentralitäts-Parameter	Beobachtete Schärfe <sup>a</sup>
Aufgabentyp	Sphärität angenommen	18.152	3	6.051	4.177	.007	.073	12.532	.848
	Greenhouse-Geisser	18.152	2.241	8.101	4.177	.014	.073	9.359	.760
	Huynh-Feldt	18.152	2.388	7.600	4.177	.012	.073	9.977	.780
	Untergrenze	18.152	1.000	18.152	4.177	.046	.073	4.177	.519
Aufgabentyp * Ausbildung	Sphärität angenommen	20.515	3	6.838	4.721	.003	.082	14.164	.892
	Greenhouse-Geisser	20.515	2.241	9.156	4.721	.008	.082	10.578	.813
	Huynh-Feldt	20.515	2.388	8.590	4.721	.007	.082	11.276	.832
	Untergrenze	20.515	1.000	20.515	4.721	.034	.082	4.721	.569
Fehler(Aufgabentyp)	Sphärität angenommen	230.303	159	1.448					
	Greenhouse-Geisser	230.303	118.751	1.939					
	Huynh-Feldt	230.303	126.583	1.819					
	Untergrenze	230.303	53.000	4.345					

a. Unter Verwendung von Alpha = .05 berechnet

Tests der Zwischensubjekteffekte									
Maß: Präferenz									
Transformierte Variable: Mittel									
Quelle		Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat	Nichtzentralitäts-Parameter	Beobachtete Schärfe <sup>a</sup>
Konstanter Term		384.048	1	384.048	679.687	.000	.928	679.687	1.000
Ausbildung		8.512	1	8.512	15.065	.000	.221	15.065	.968
Fehler		29.947	53	.565					

a. Unter Verwendung von Alpha = .05 berechnet

## F PERSÖNLICHE ERKLÄRUNG

---

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbständig verfasst und die den verwendeten Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe

Zürich, 30. September 2015

---

Oliver Deseö