

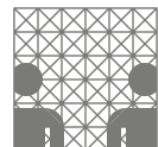


# Verbale Assistenz für virtuelle taktile Karten. Effiziente, multi-modale Interfaces für sehbehinderte Menschen.

Kris Lohmann

lohmann@informatik.universität-hamburg.de  
Supervisors: Christopher Habel & Carola Eschenbach  
WSV / Fachbereich Informatik  
Universität Hamburg

Oberseminar Mensch-Computer-Interaktion  
Veranstalter: Martin Christof Kindsmüller



## Ziel und Motivation

Verbesserung von Karten für blinde und sehbehinderte Menschen durch die Unterstützung von virtuellen taktilen Karten mit sprachlicher Assistenz



(Bild von OpenStreetMap - Veröffentlicht unter CC-BY-SA 2.0)

## Motivation & Ziel

- Karten: wichtige externe Repräsentationen
- Visuelle Karten unzugänglich für *blinde und sehbehinderte* Menschen
- Nachteil: Sequentielle Exploration erhöht Integrationsanforderungen und verringert Geschwindigkeit und Genauigkeit  
(Loomis et al. 1991, Röder et al. 2001)
- Multi-Modalität kann Benutzbarkeit von taktilen Karten erhöhen (Jacobson 2002, Wang et al. 2009)

# Forschungsstand multi-modale Karten

- Forschung an multi-modalen Karten
  - NOMAD (Parkes, 1988)
  - Brailledis 9000 Audio-Haptic Browser (Zeng & Weber, 2010)
  - OMERO (De Felice et al. 2007)
  - ...
- *Aber: Die Schnittstellen nutzen Effektivität räumlicher Sprache nicht voll aus!!!*

→ Das ist noch zu beweisen

# Forschungsstand verbale Beschreibungen

- Lernen sehende Menschen mit verbundenen Augen Räume besser, wenn sie sprachliche (Kurz-) Beschreibungen kriegen? (Giudice, Bakdash, und Legge 2007)
  - Ja!
- Wiederholung mit blinden Teilnehmern (Giudice 2004)
  - Vergleichbare Ergebnisse

# VAVETaM Projekt

- **V**erbally **A**ssisting **V**irtual **E**nvironment **T**actile **M**aps  
(Lohmann, Habel, Kerzel, 2010)
- Ziel: VAVETaM-System, Teilbereiche
  - Sprache
  - Eventerkennung in der Haptik (Matthias Kerzel)
  - Benutzermodell

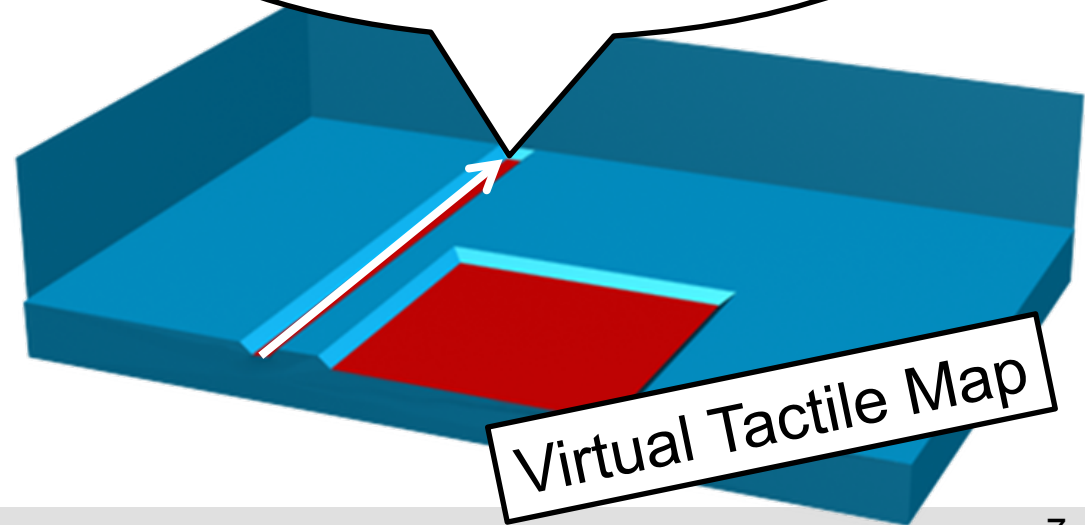
# Einleitung: Virtual Tactile Maps & Assistenz

PHANToM Omni



Sprachliche Assistenz

Das ist die Vogt-Kölln-Straße, rechts daneben ist das Informatikum.



Virtual Tactile Map

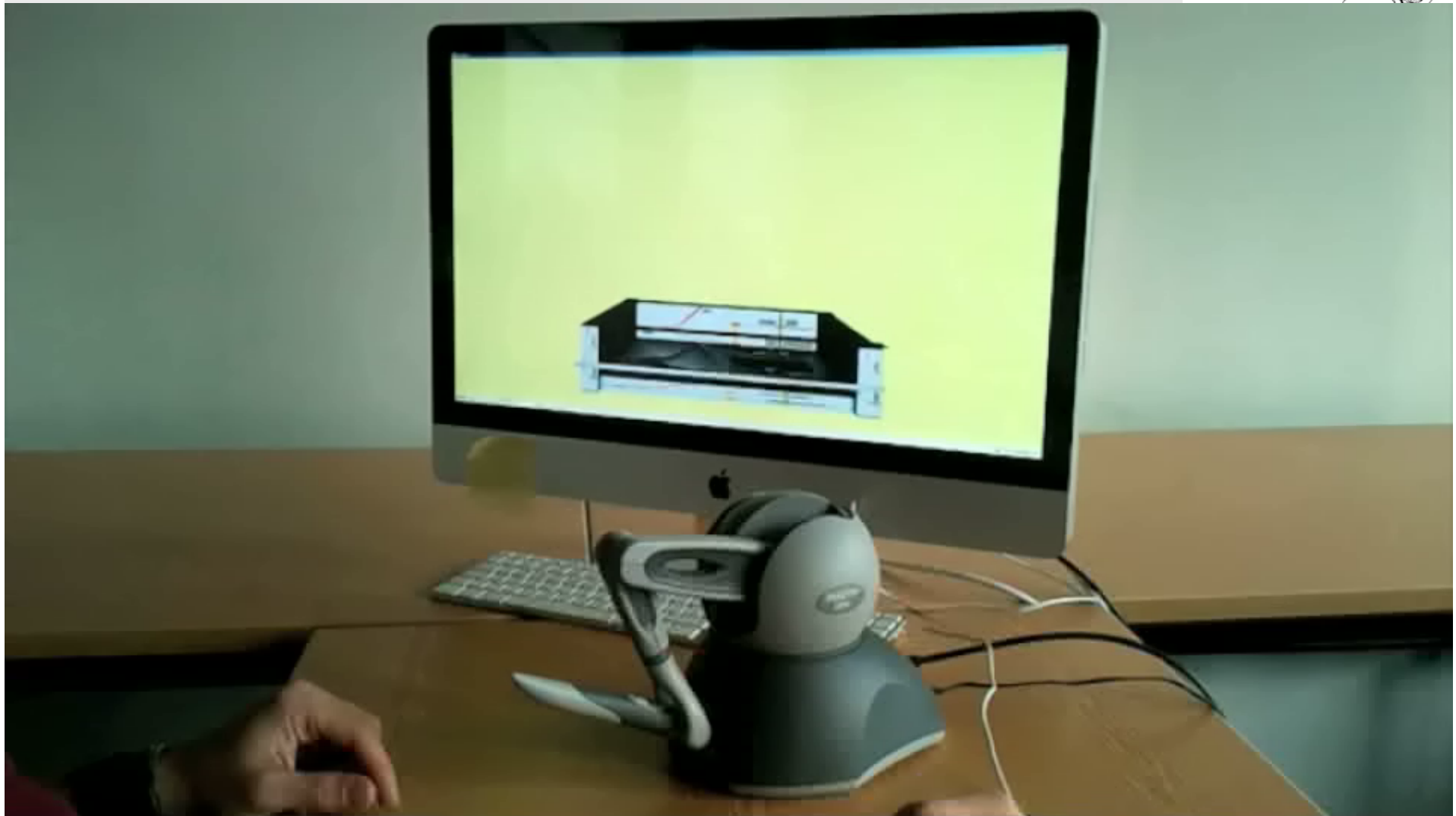
# Technisches

- Grundlage für *virtuelle* Karten:  
Sensable Phantom  
Schnittstelle für virtuelle Haptik
- Vorteile virtueller Karten
  - Verfügbarkeit
  - Aufwand
  - Flexibilität
- Nachteil
  - 1-Punkt-Interaktion



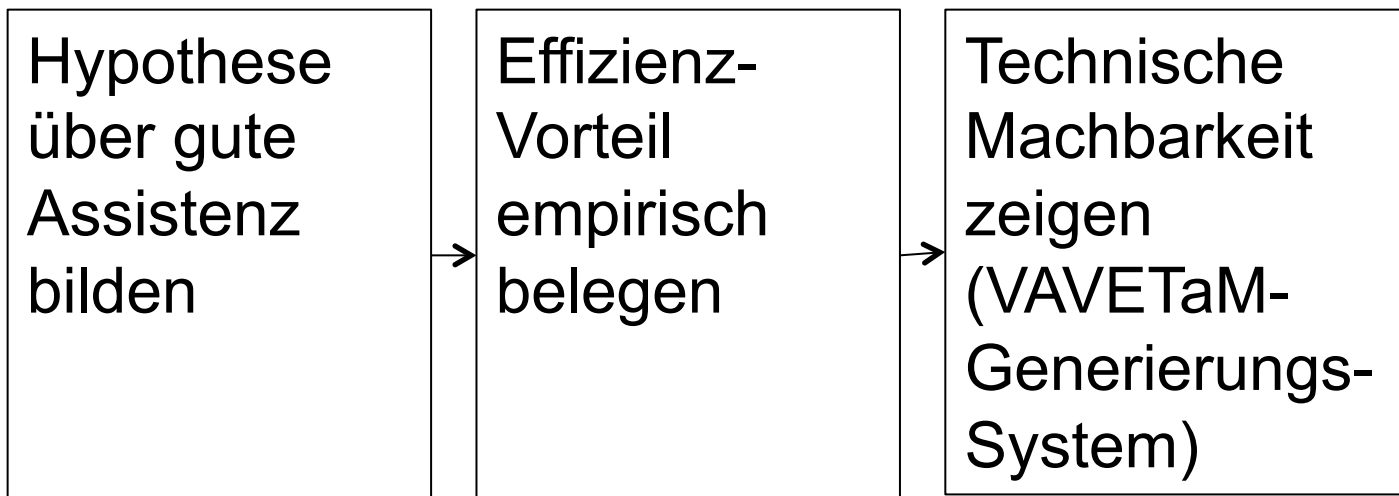
(Bild von OpenStreetMap –  
Veröffentlicht unter CC-BY-SA 2.0)





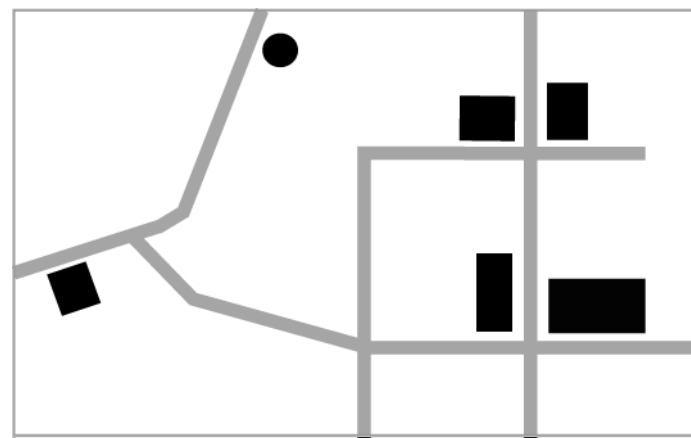
(Video, Phantom Omni und Kartenmodell)

# Forschungsschritte zur Entwicklung der VAVETaM-Sprache



# Exkurs: Terminologie

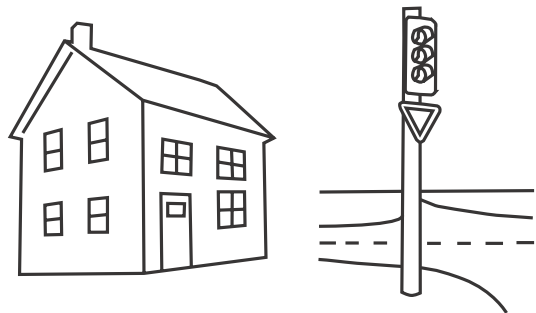
- **Tracks:** Strukturen, die Lokomotion ermöglichen, z. B. Hauptstraßen und Feldwege (~Path bei Lynch, 1960)
- **Landmarken:** potentielle Landmarken, in unseren Karten Bäume und Gebäude (aber unter den Begriff fällt natürlich mehr)



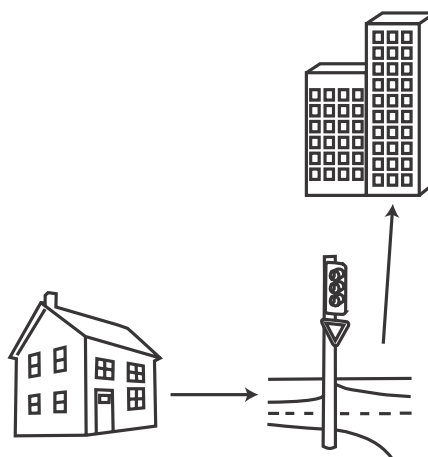
## Terminologie (cont.)

- **Effizienz:** Geschwindigkeit und Genauigkeit des Wissenserwerbs
- Das heißt, eine Schnittstelle ist effizienter, wenn ...
  - mehr Wissen in der gleichen Zeit erworben wird oder
  - gleich viel Wissen in kürzerer Zeit erworben wird

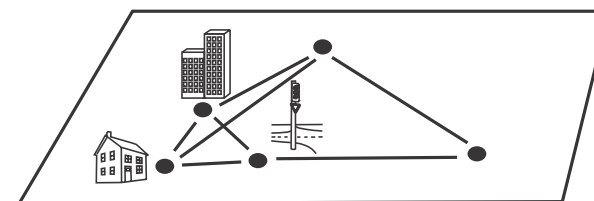
# Typen räumlichen Wissens



Landmark-  
Wissen



Routen-  
Wissen



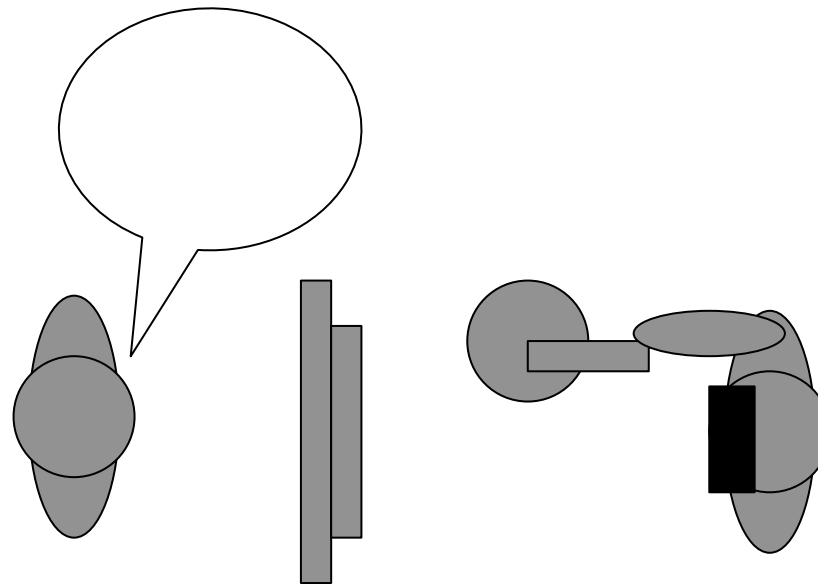
Überblicks-  
Wissen

# Gliederung

1. Motivation & Technik
2. Was gesagt werden soll
3. Experiment 1: Effekt verbaler Assistenz bei Sehenden
4. Experiment 2: Effekt verbaler Assistenz bei blinden und sehbehinderten Menschen
5. VAVETaM-Prototyp
6. Ausblick und Zusammenfassung

# Assistenz-Korpus

- Sprachliche Assistenz, die Menschen Karten-Explorierern mit verbundenen Augen geben
- Eine erste Idee, was gesagt werden könnte



## Korpus: Beispiele

- “Das ist die neue Gasse, die rechts in einer Sackgasse endet und links ebenfalls in einer Sackgasse endet“
- “Die Weberstraße ist parallel zur Faberstraße und zur Markstraße und liegt zwischen den beiden“
- “Auf dem Bahnhofplatz liegt noch ein weiteres Gebäude und drei Bäume”
- “Das ist die Sackgasse, in die die Hochstraße auf der linken Seite mündet”

(Lohmann, Eschenbach & Habel 2011)



# Message Classes

- Gruppierung der Information im Korpus nach Informationen
- Beispiele für Tracks
  - Junctions
  - Extension
  - Spatial Relations

# Gliederung

1. Motivation & Technik
2. Was gesagt werden soll
3. Experiment 1: Effekt verbaler Assistenz bei Sehenden
4. Experiment 2: Effekt verbaler Assistenz bei blinden und sehbehinderten Menschen
5. VAVETaM-Prototyp
6. Ausblick und Zusammenfassung

# Experiment 1: Assistenz-Effekt auf Teilnehmer mit verbundenen Augen

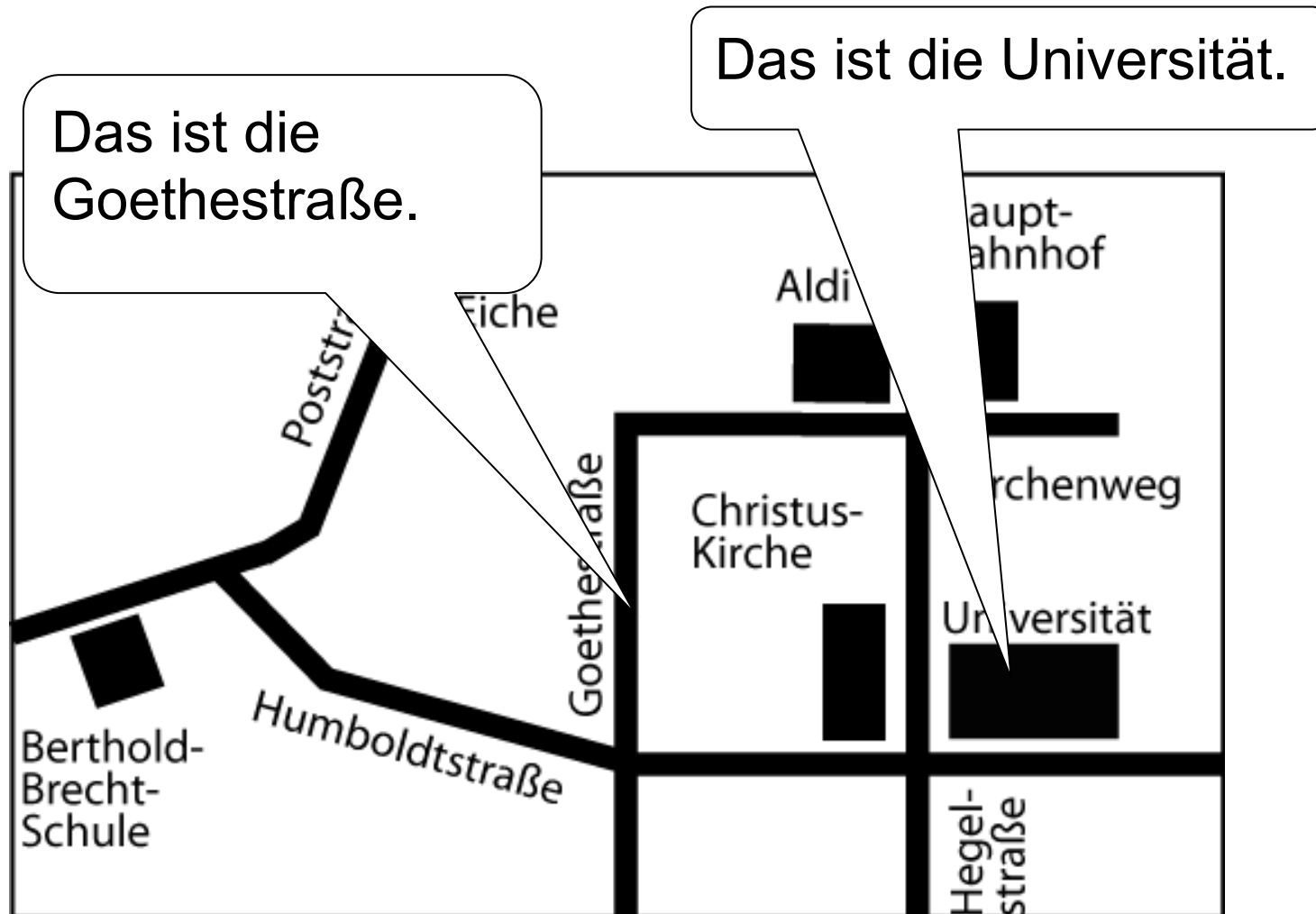
- *Hypothese*
  - Sprachassistenz (mit dem Assistenzset) ist hilfreich, daher Wissen bei mehr Assistenz

Anmerkung (für Online-Version): Das Experiment wird diskutiert in Lohmann und Habel (2012, in press):  
Extended Verbal Assistance Facilitates Knowledge Acquisition of Virtual Tactile Maps. Accepted for Publication in *Proceedings of Spatial Cognition 2012*

# Methode und Ablauf

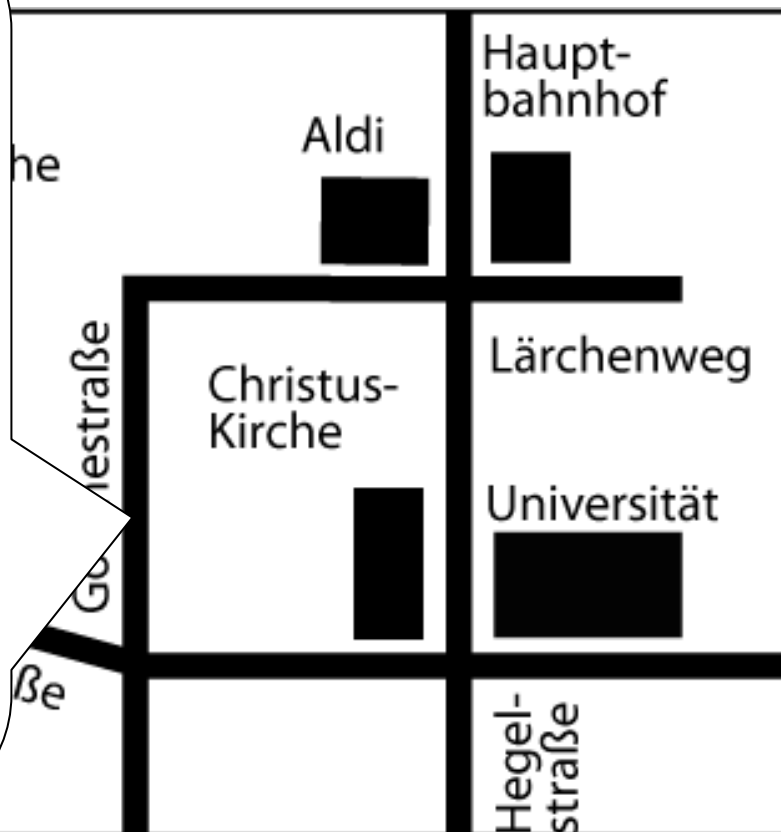
- Teilnehmer
  - 24 ausgewertete Teilnehmer (14 männlich, Altersdurchschnitt 24.7 Jahre,  $SD = 3.3$ )
- Material und Methode
  - Repeated-Measures-Design
  - Wizard-of-Oz-Experiment
  - Kartenlernen unter zwei Bedingungen
  - 8 min Lernzeit
    - **Simple assistance**
    - **Extended assistance**

# “Simple Assistance”-Bedingung

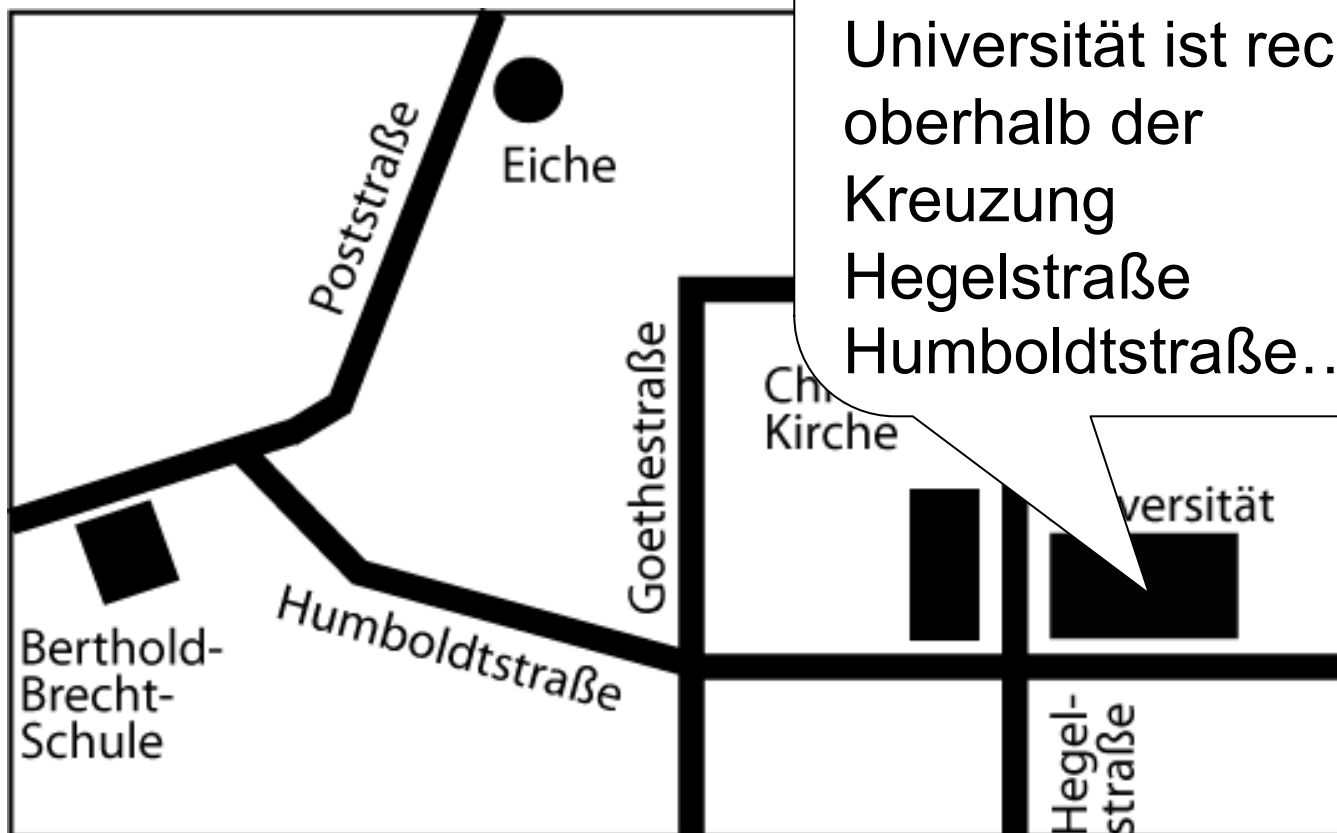


# “Extended Assistance”-Bedingung

Das ist die Goethestraße. Die Goethestraße ist parallel zur Hegelstraße. Die Goethestraße wird unten durch den Kartenrand begrenzt und endet oben in einer Ecke mit dem Lärchenweg...

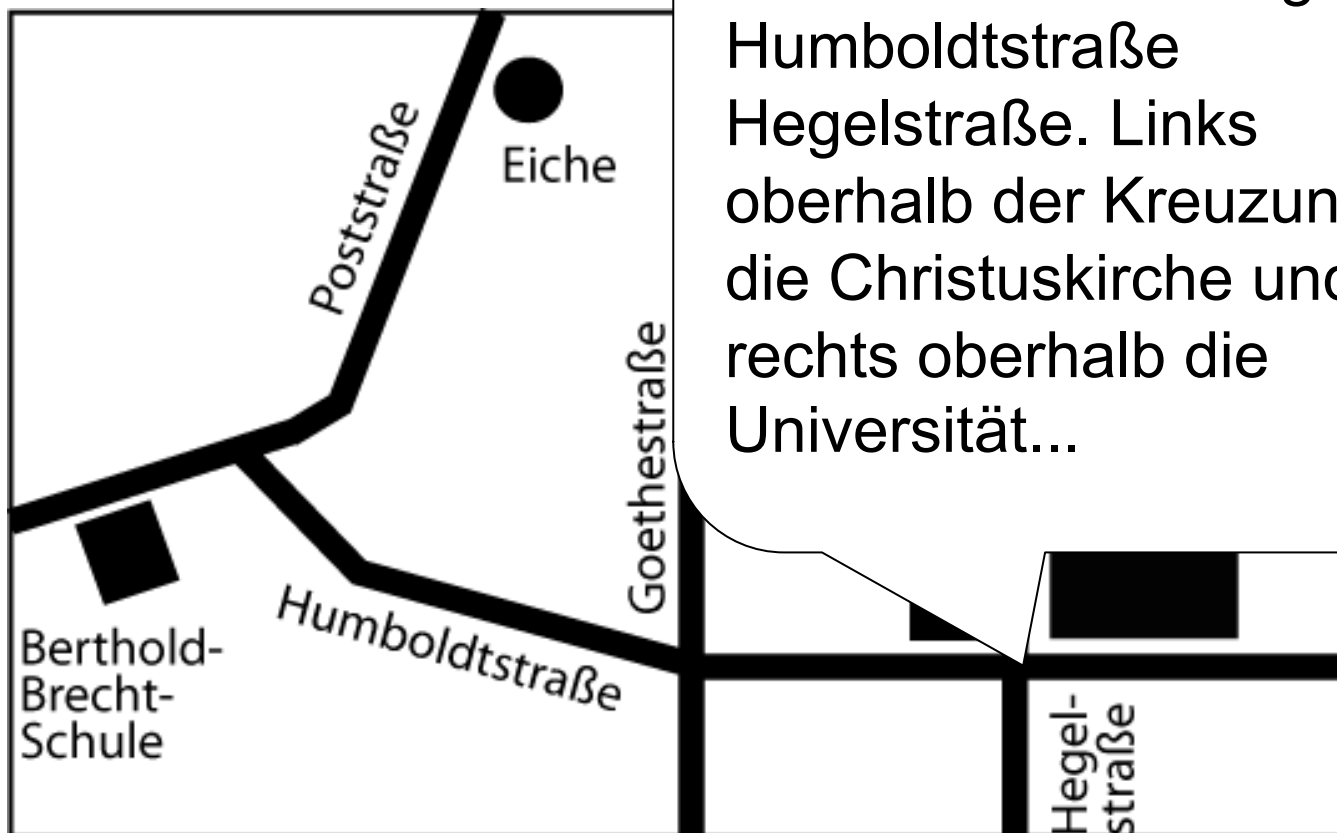


# “Extended Assistance”-Bedingung



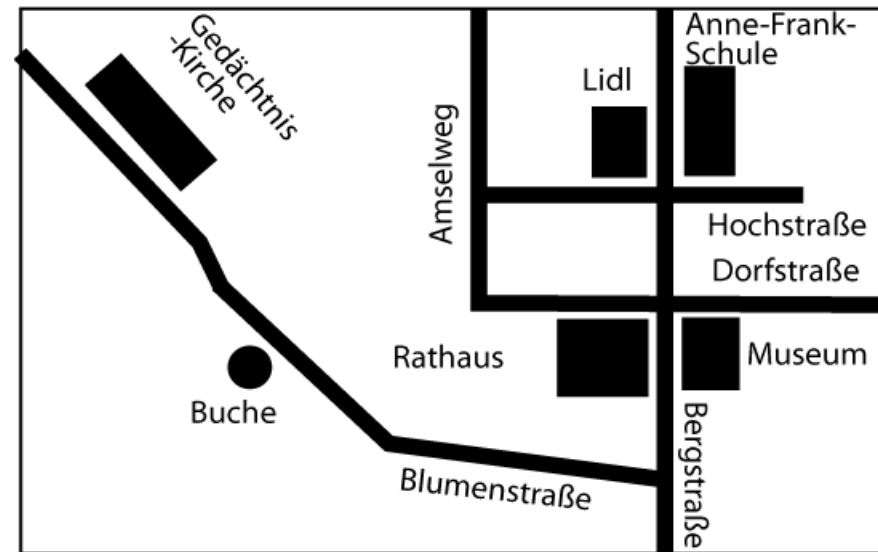
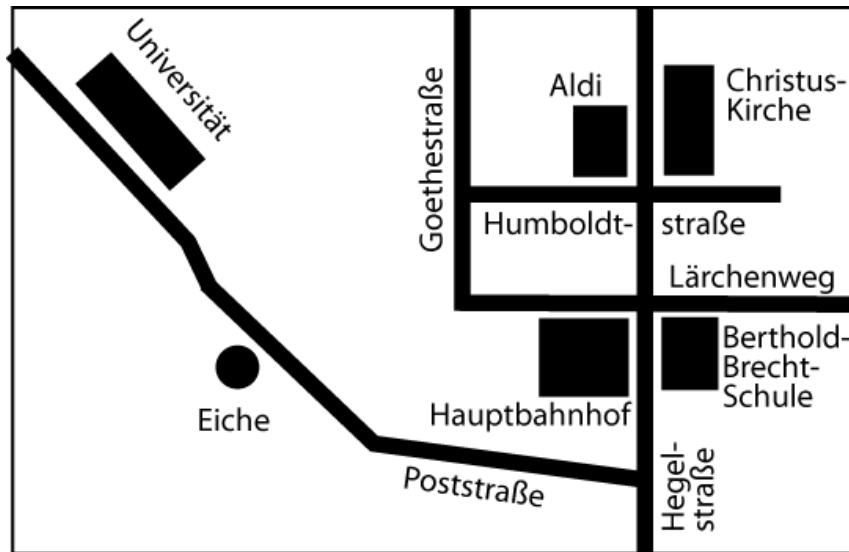
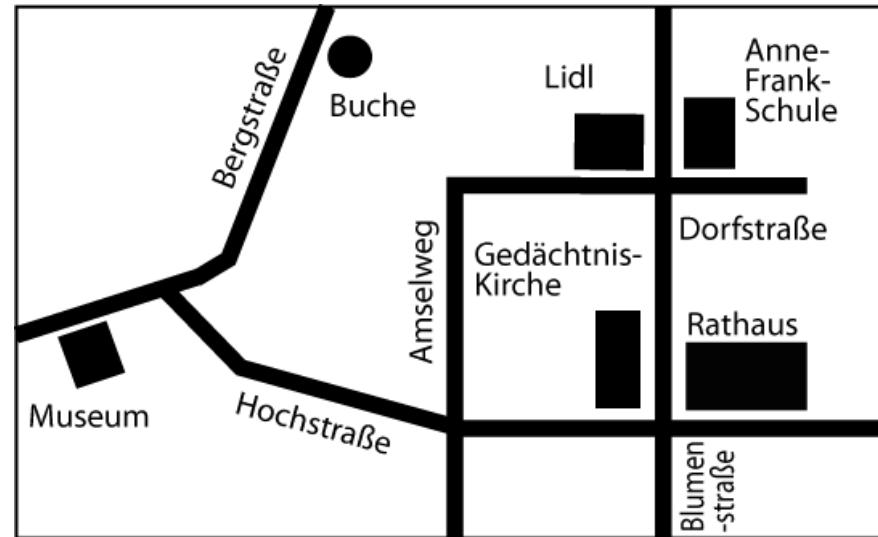
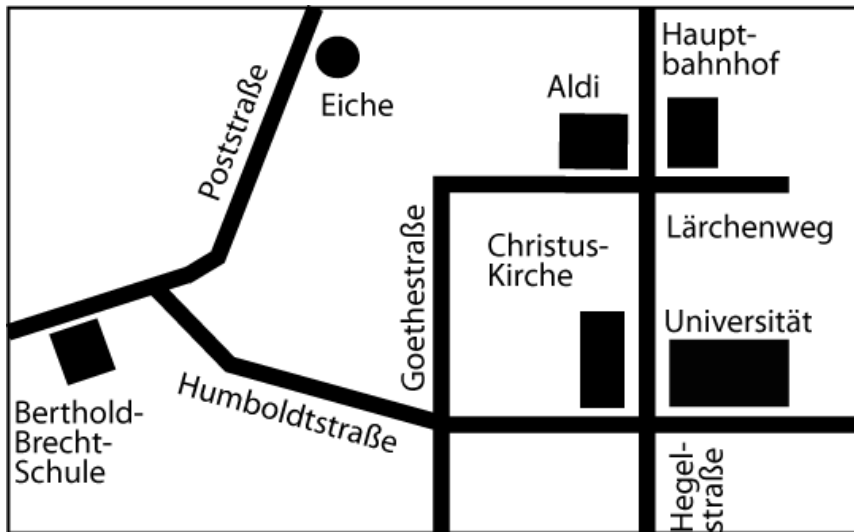
Das ist die  
Universität. Die  
Universität ist rechts  
oberhalb der  
Kreuzung  
Hegelstraße  
Humboldtstraße...

# “Extended Assistance”-Bedingung

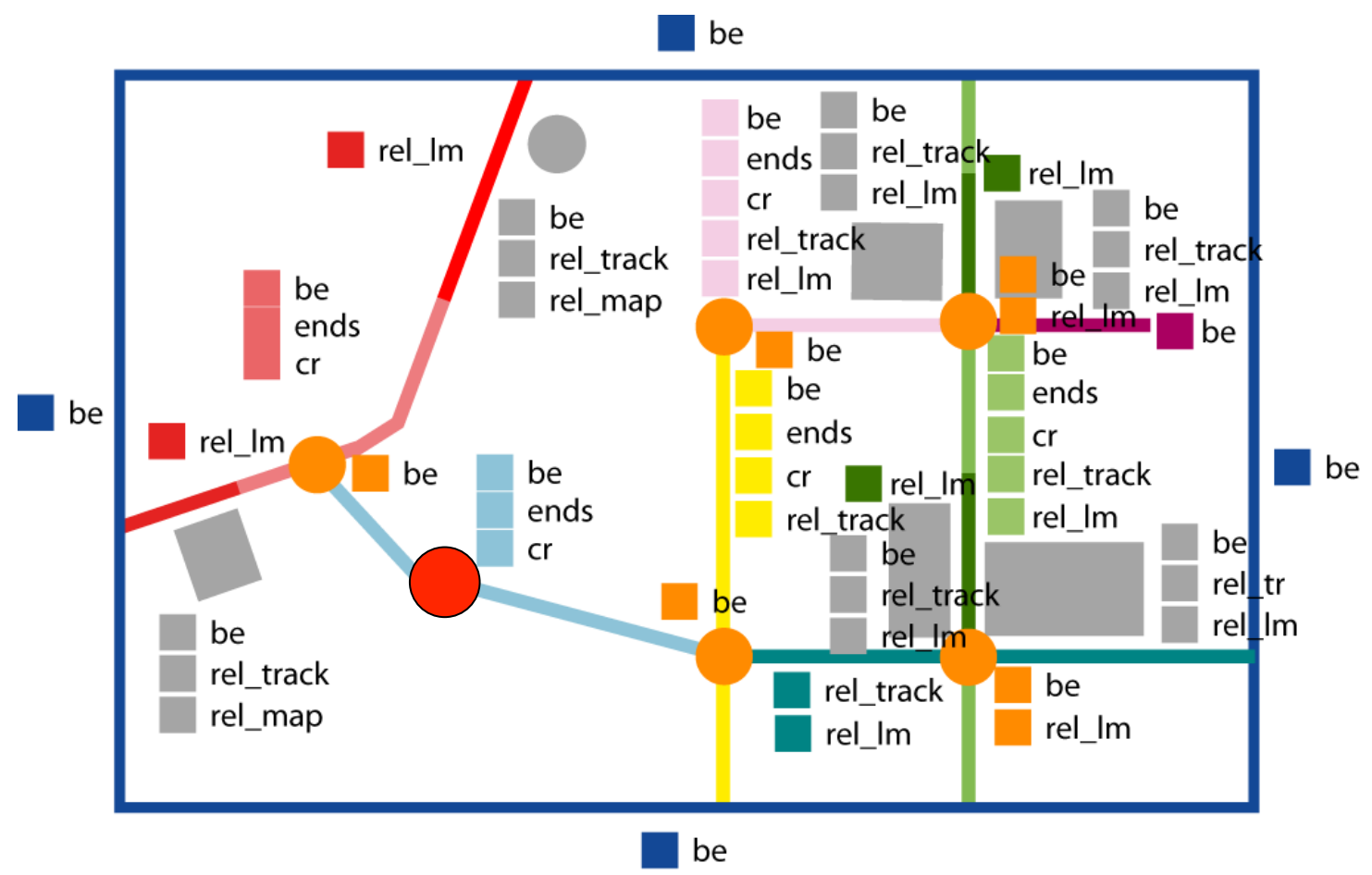


Das ist die Kreuzung  
Humboldtstraße  
Hegelstraße. Links  
oberhalb der Kreuzung ist  
die Christuskirche und  
rechts oberhalb die  
Universität...

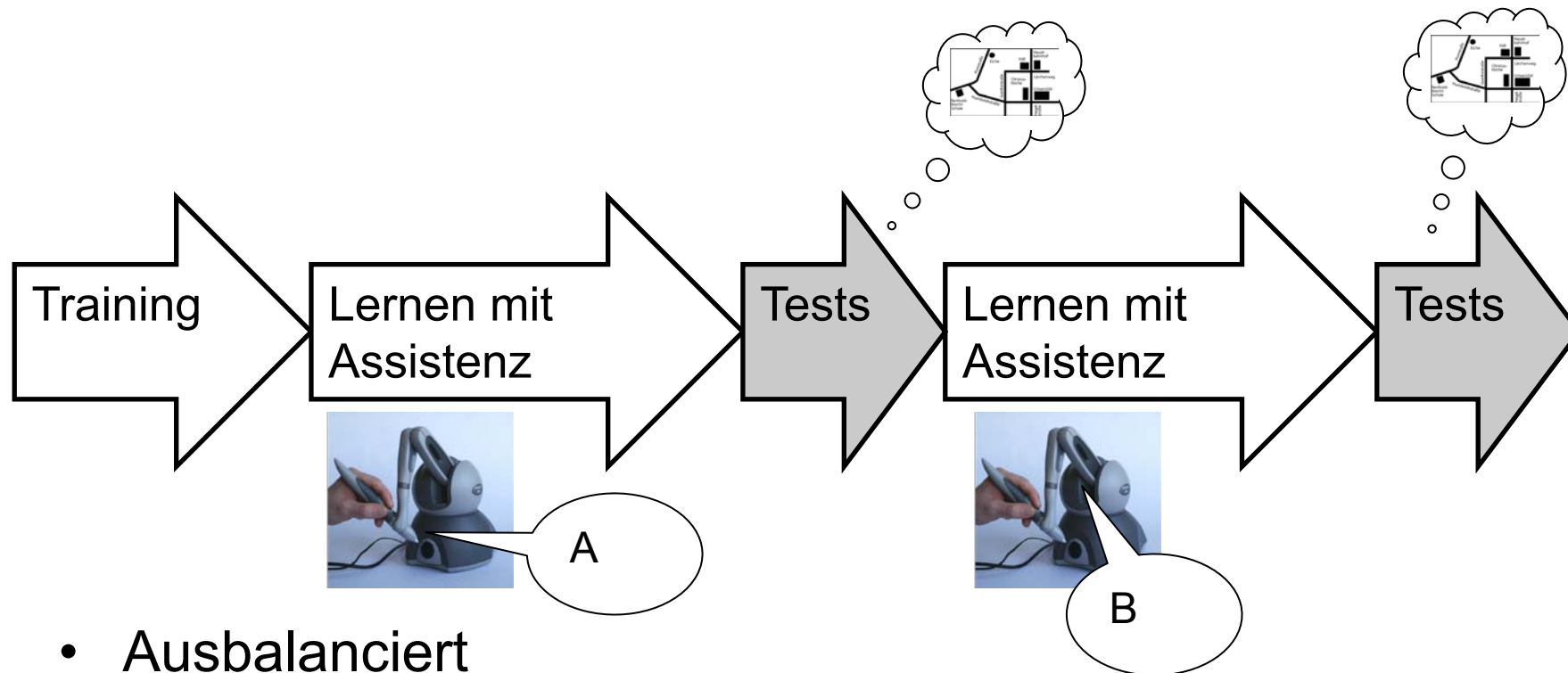




# Wizard-of-Oz Software



# Ablauf



- Ausbalanciert
  - Namensset/Bedingung
  - Karte/Bedingung
  - Reihenfolge der Bedingungen

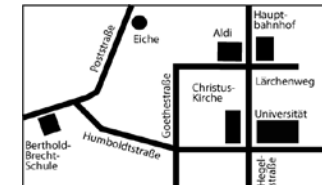
# Tests räumlichen Überblickswissens

- 3 Tests
  - 20 Fragen zu Relationen (Relation Questions)
  - Kartenskizzen
  - Recognition Puzzle

# Relation Questions

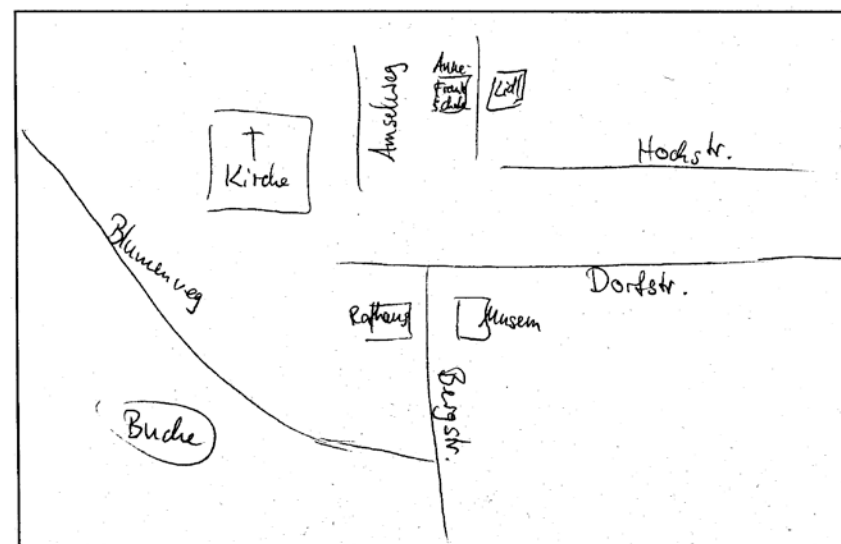
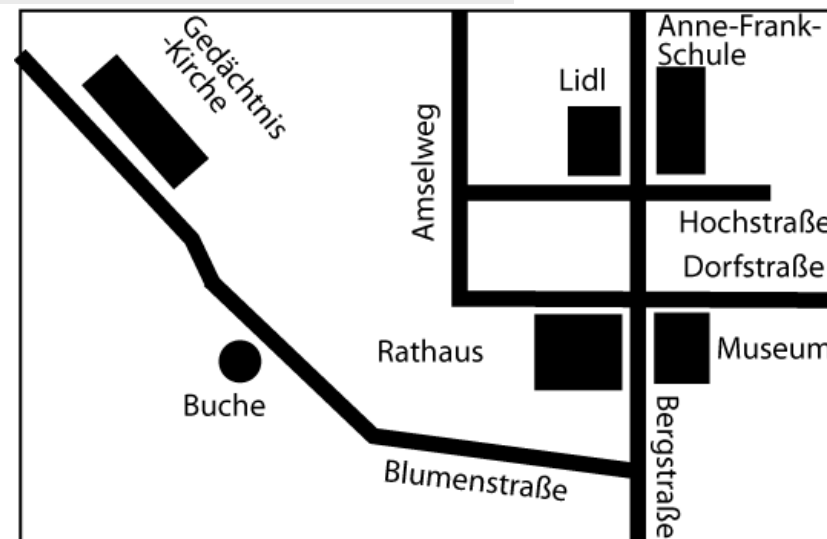


- Antwortoptionen: “Ja”, “Nein”, “Ich weiß nicht”
- Randomisierte Reihenfolge
- 10 Fragen zu Tracks
  - Beispiel: “Mündet die Humboldtstraße in die Poststraße ein?”
  - 5 richtig beantwortet mit “Ja”, 5 mit “Nein”
- 10 Fragen zu Landmarken
  - Beispiel: “Ist die Christuskirche unterm Aldi?”
  - 5 richtig beantwortet mit “Ja”, 5 mit “Nein”
- 4 zusammenpassende Fragesets

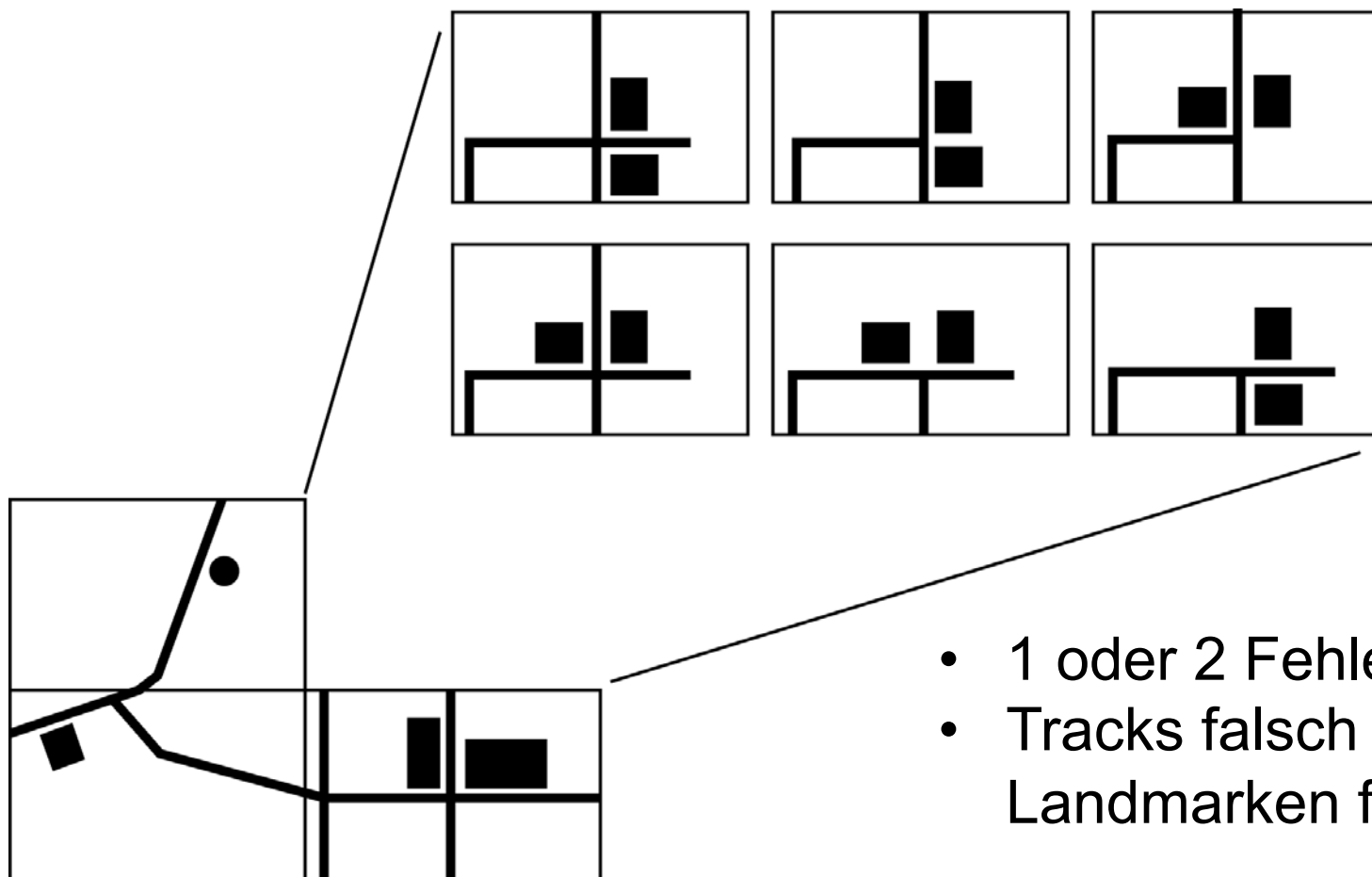


# Sketch Maps

- Der Rahmen der Karte war vorgegeben
- Instruktion: 'Zeichne alles auf, an das Du Dich erinnerst'

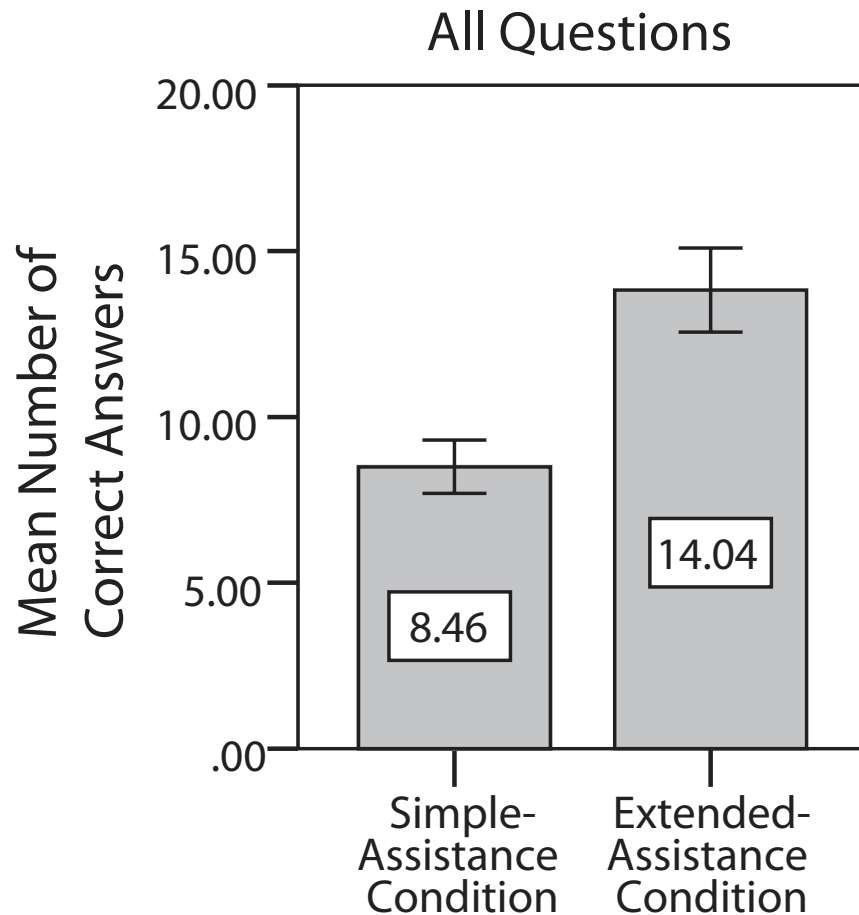


# Recognition Puzzle



- 1 oder 2 Fehler pro Teil
- Tracks falsch und/oder Landmarken falsch

# Resultate: Relation Questions



Fehlerbalken: 95 % CI

- Insgesamt signifikant mehr richtig beantwortete Fragen ( $t(23) = 8.08, p < .001, r = .86$ )



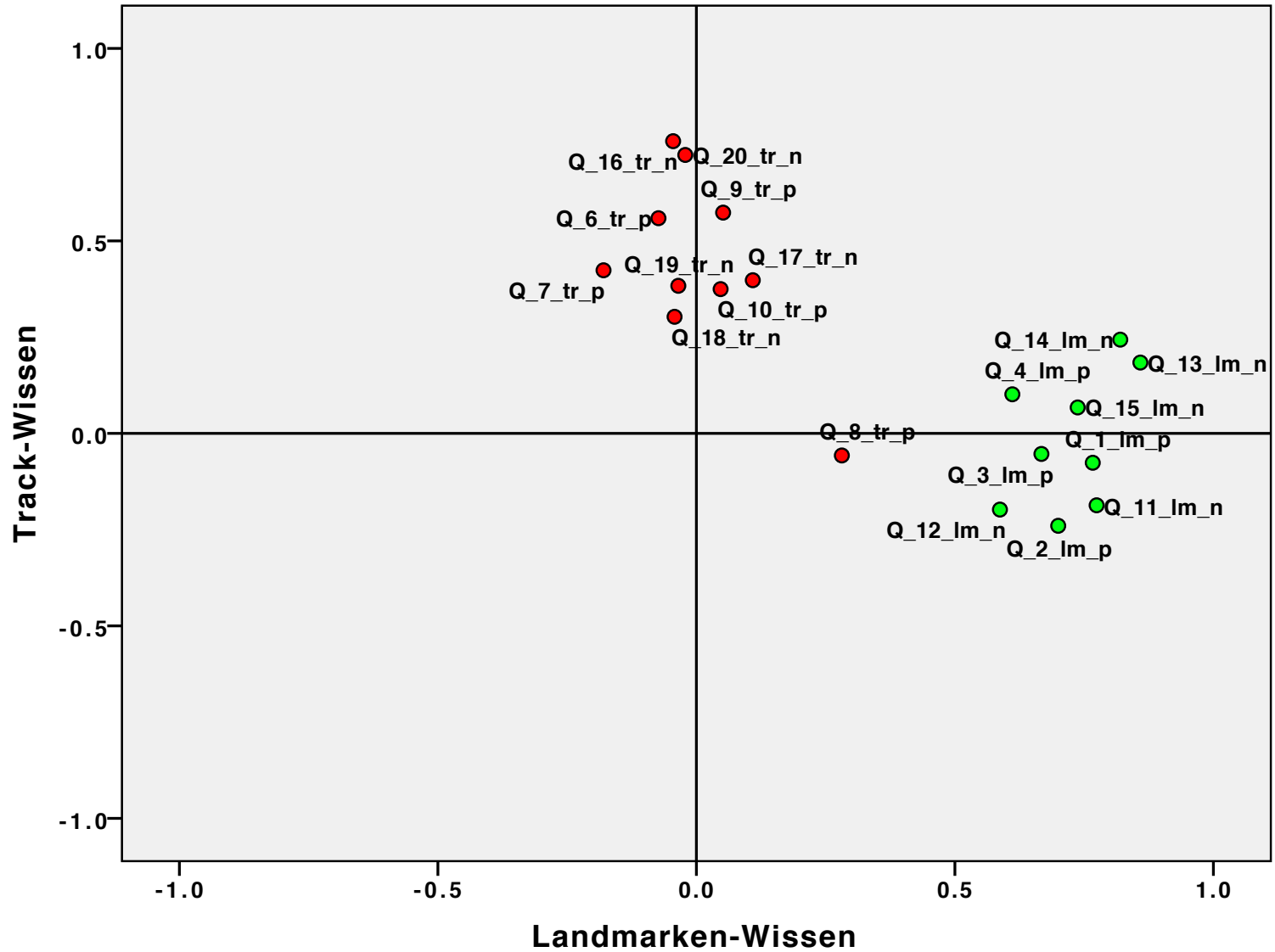
## Resultate Relation Questions: 2 Komponenten

- Bedingung hat Effekt auf einige Fragen
- Track-Fragen anders als Fragen zu Landmarken
- Bestätigt in Hauptkomponenten-Analyse (PCA)  
→ bekannt?

# Hauptkomponentenanalyse (PCA)

- Ein Verfahren, mit dem nicht direkt gemessene Variablen identifiziert werden können
- Auf der Basis von Korrelationen werden zusammenhängende Variablen gesucht (Bortz und Schuster 2010)

# Komponentendiagramm im rotierten Raum



- Track-Fragen
- Landmarken-Fragen

# Ergebnis: Hauptkomponentenanalyse



Fragen zu Landmarken

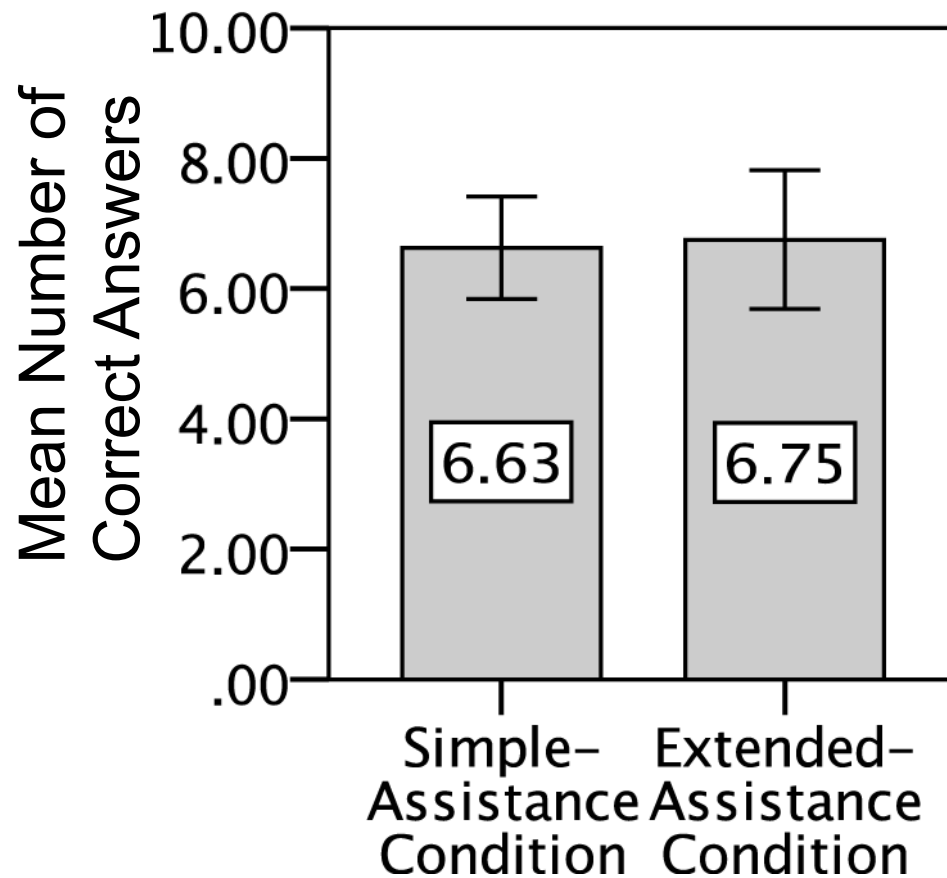
Fragen zu Tracks

Question Number	Component 1 (Landmark Knowledge)	Component 2 (Track Knowledge)
1	<b>.776</b>	.125
2	<b>.784</b>	-.038
3	<b>.669</b>	.121
4	<b>.546</b>	.246
11	<b>.832</b>	.028
12	<b>.656</b>	-.029
13	<b>.749</b>	.382
14	<b>.685</b>	.426
15	<b>.684</b>	.247
6	-.319	<b>.483</b>
7	-.362	<b>.334</b>
8	.298	.020
9	-.204	<b>.527</b>
16	-.380	<b>.669</b>
17	-.071	<b>.384</b>
18	-.175	<b>.261</b>
19	-.204	<b>.335</b>
20	-.342	<b>.643</b>

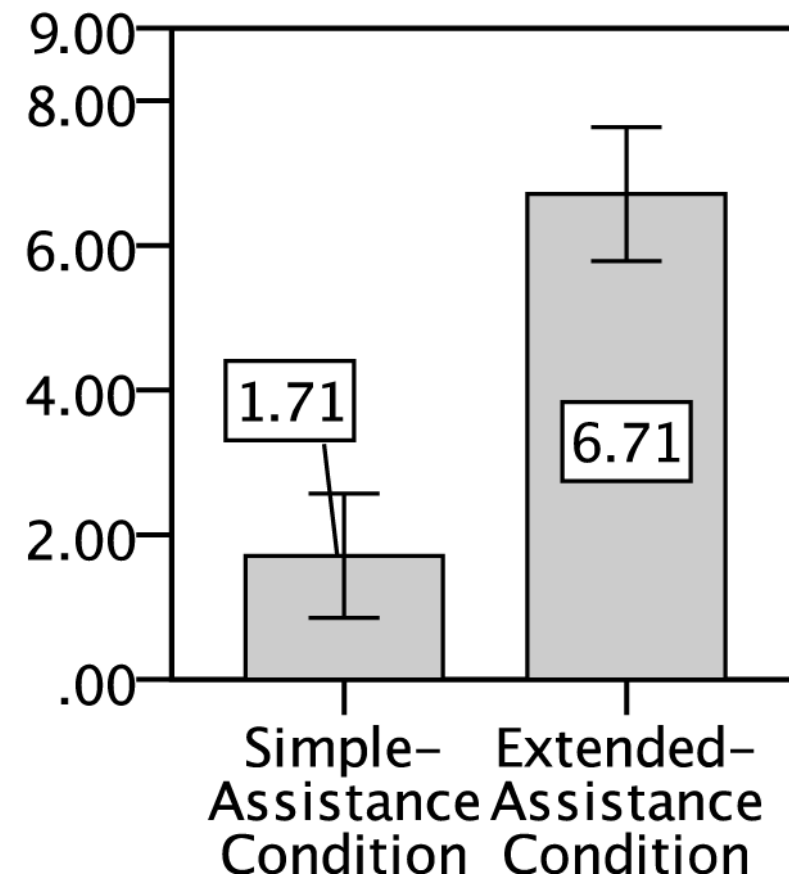
# Resultate: Relation Questions



### Track Questions



### Landmark Questions



Fehlerbalken: 95 % CI

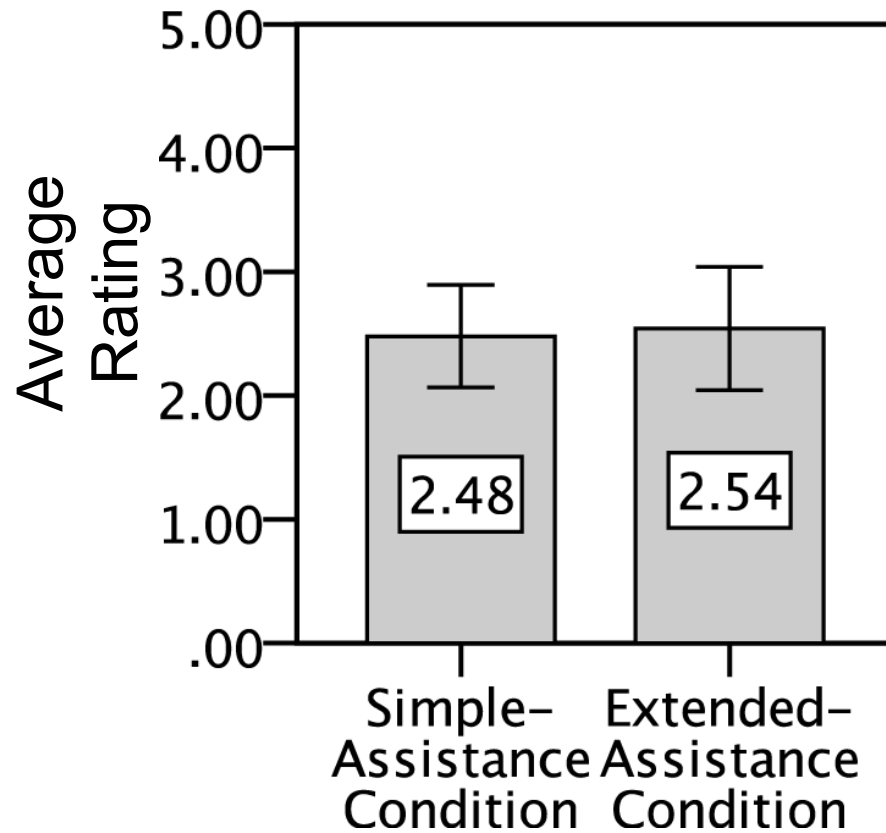
# Sketch Maps

- Bewertet von mir und einem weiteren Rater, ohne Wissen über die Assistenzbedingung, auf einer 1–5 Likert-type Skala
- In zwei Hinsichten, entsprechend den Subdimensionen: ‘Wie gut ist die Straßenkonfiguration der Originalkarte in der Skizze wiedergegeben?’ ( $ICC(3, 1) = .68$ ) und ‘Wie gut ist die Konfiguration der Landmarken wiedergegeben?’ ( $ICC(3, 1) = .72$ )

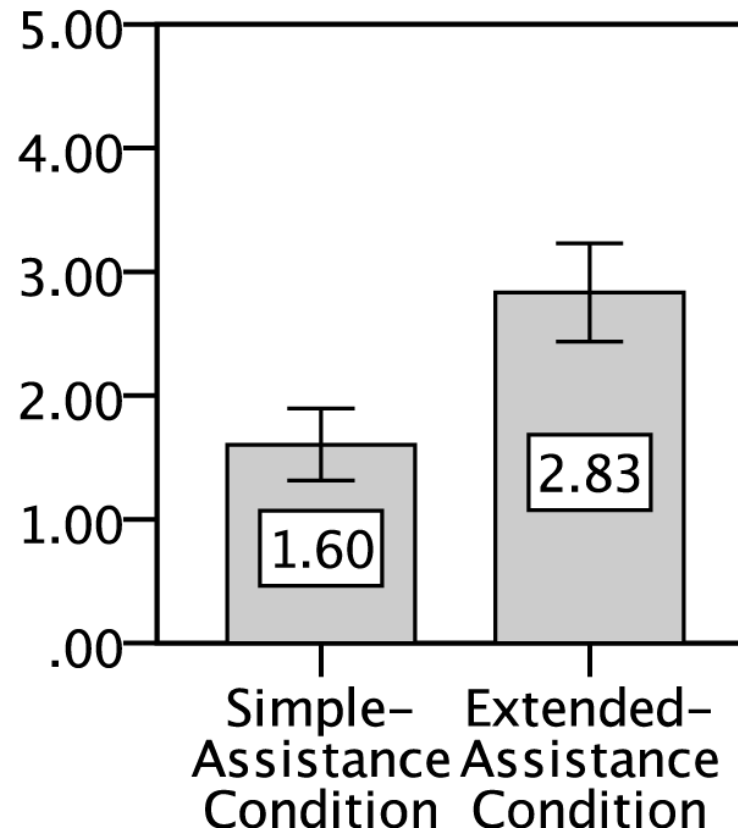
# Resultate: Sketch Maps



### Mean Track-Structure Rating

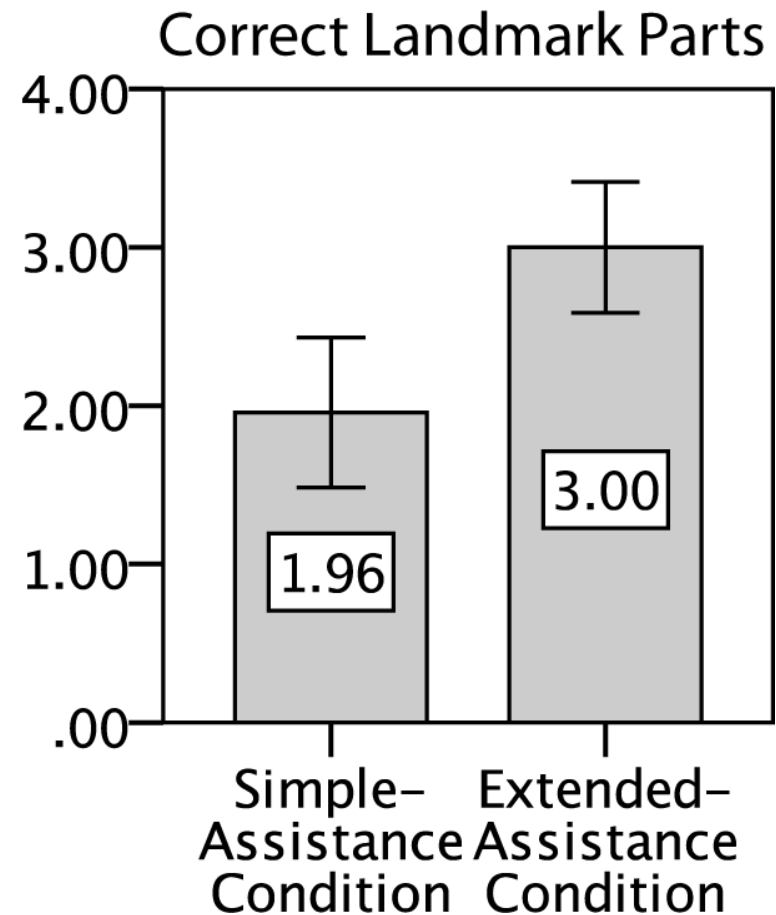
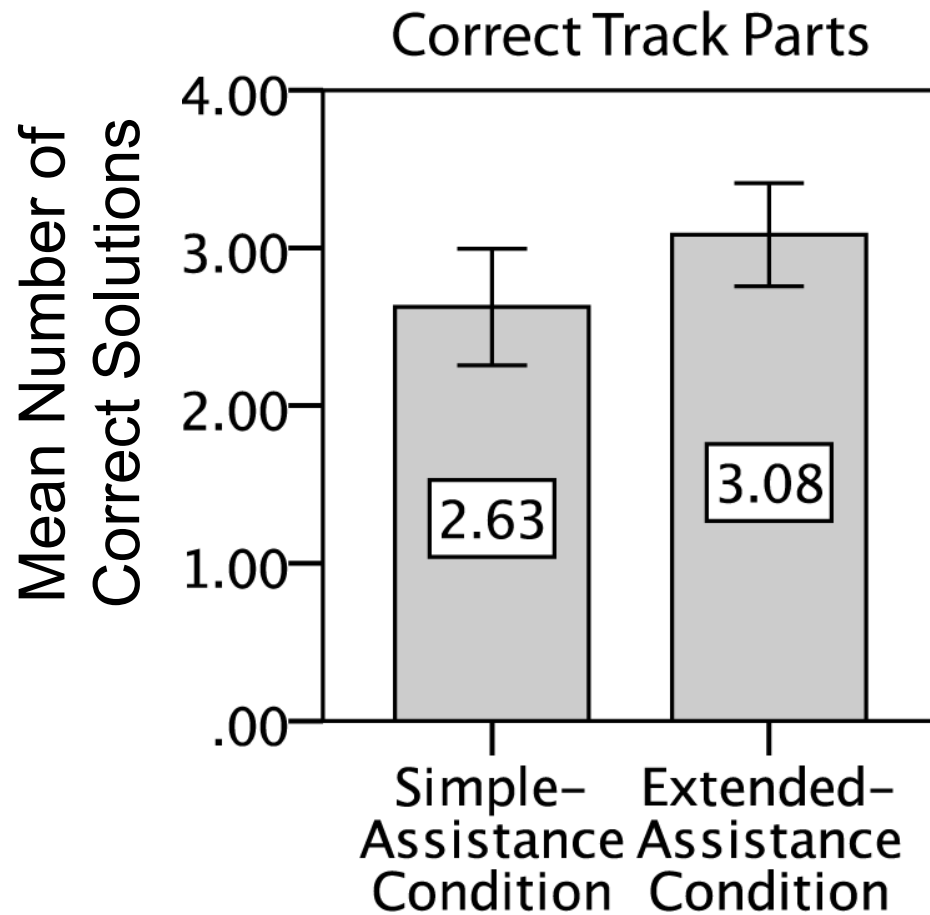


### Mean Landmark-Configuration Rating



Fehlerbalken: 95 % CI

# Resultate: Recognition Puzzle



Fehlerbalken: 95 % CI



# Experiment 1: Diskussion

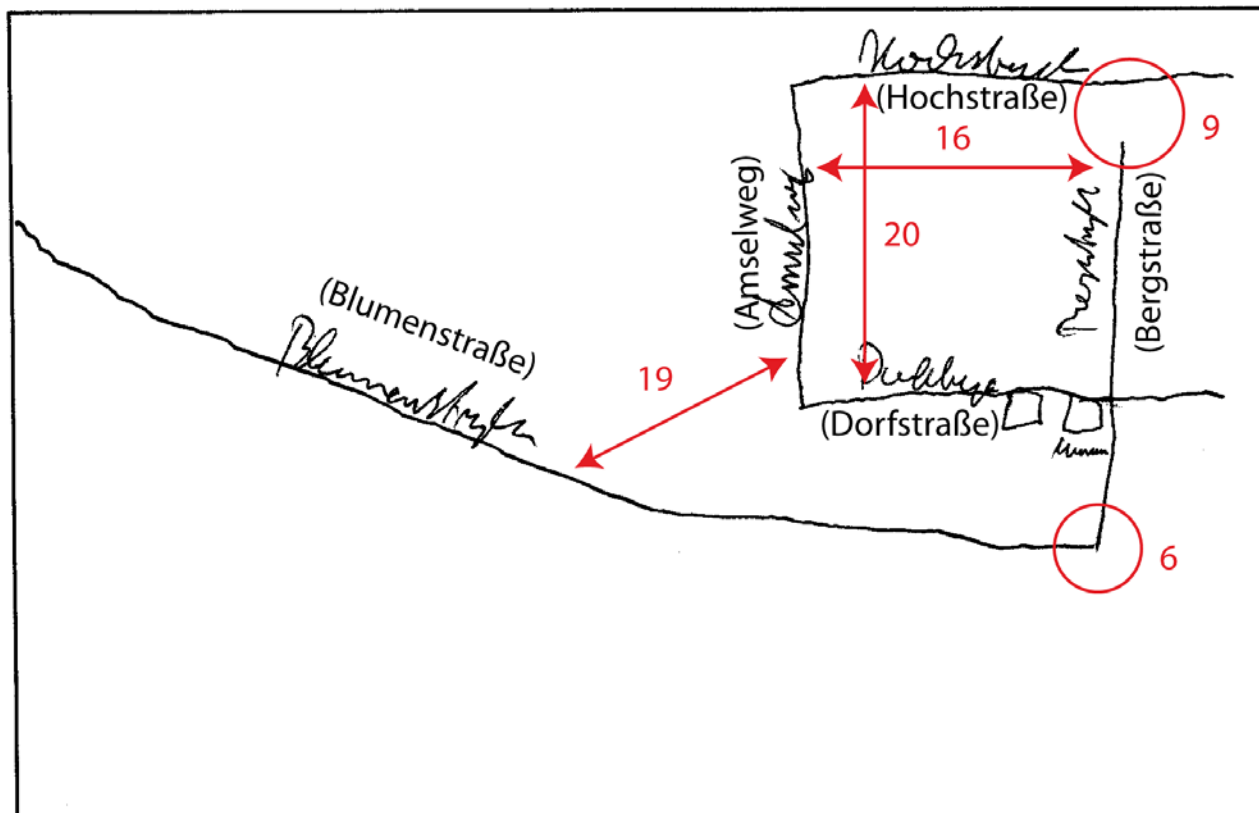
- Mixed-model ANOVA, keine signifikanten Effekte von Karte, Händigkeit, Reihenfolge der Bedingungen, Geschlecht (außer beim Recognition Puzzle) oder Namensset
- Alle Tests zeigen einen signifikanten ( $p < .05$ ) Effekt der Assistenzbedingung, wenn es um Landmarken-Wissen geht, ...
- ... kein Test einen signifikanten Effekt, wenn es um Track-Wissen geht
- Paired-sample t-test for equivalence zeigt, dass Straßenwissen der Relation Questions und der Skizzen statistisch equivalent ist (Wellek, 2003)



# Exkurs: Sketch Maps als Evaluationsbasis?



- Vergleich der Antworten der Relation-Questions mit dem, was auf den Skizzen dargestellt ist (Lohmann, 2011)



- 87.70 %  
Übereinstimmung  
→ Skizzen  
reflektieren  
Wissen ganz gut

?

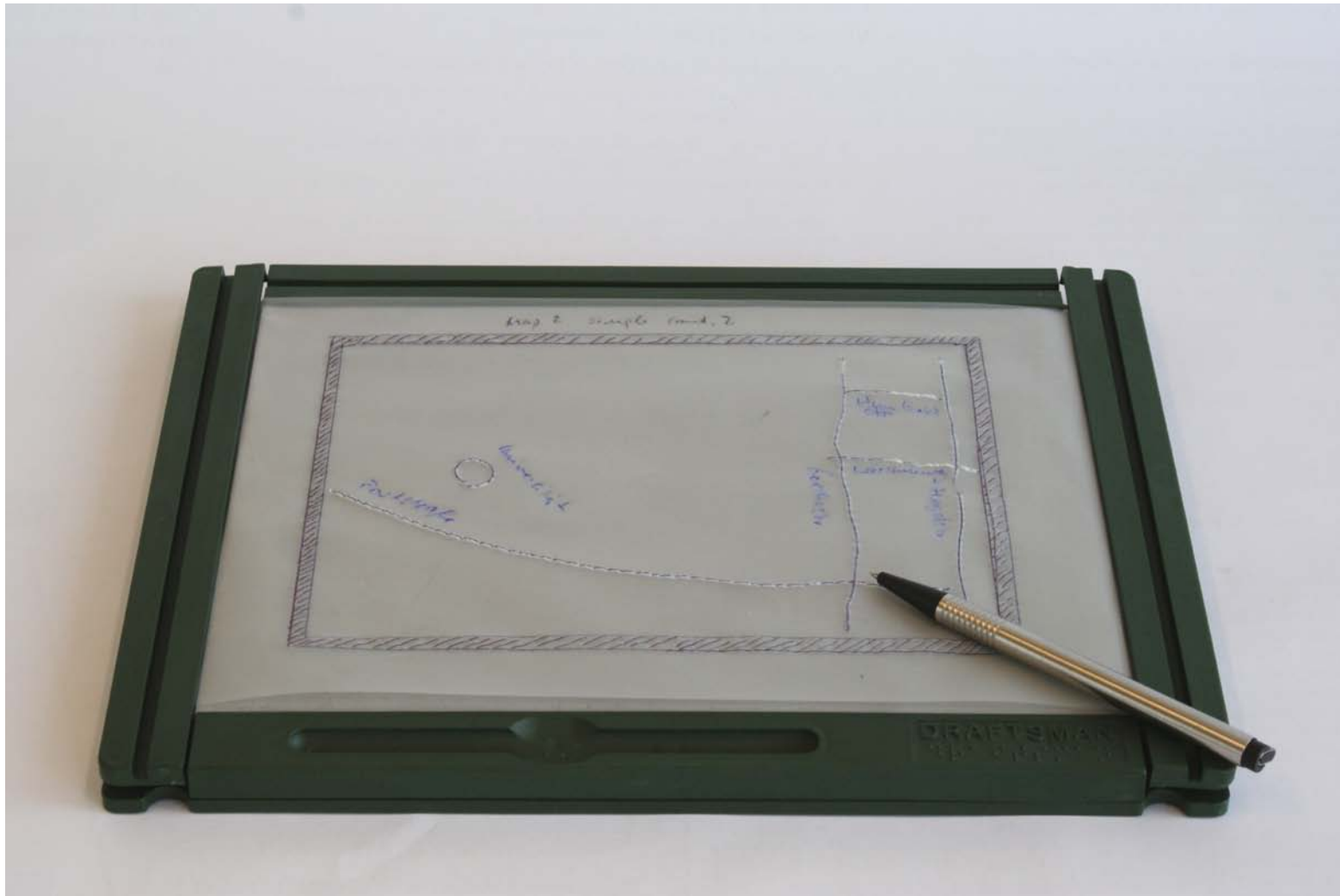
# Gliederung

1. Motivation & Technik
2. Was gesagt werden soll
3. Experiment 1: Effekt verbaler Assistenz bei Sehenden
4. Experiment 2: Effekt verbaler Assistenz bei blinden und sehbehinderten Menschen
5. VAVETaM-Prototyp
6. Ausblick und Zusammenfassung

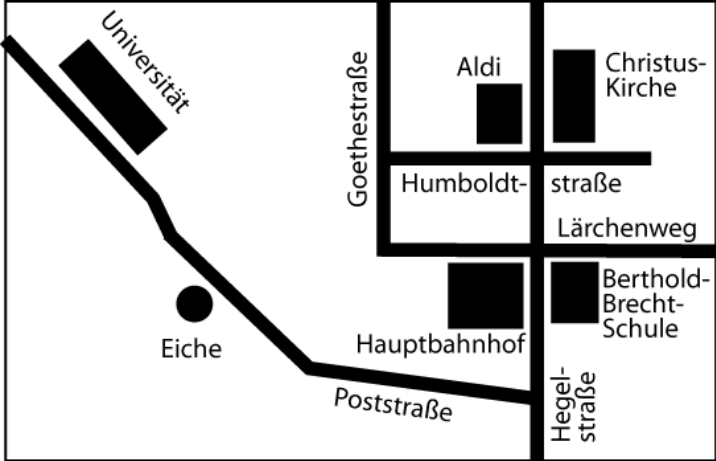
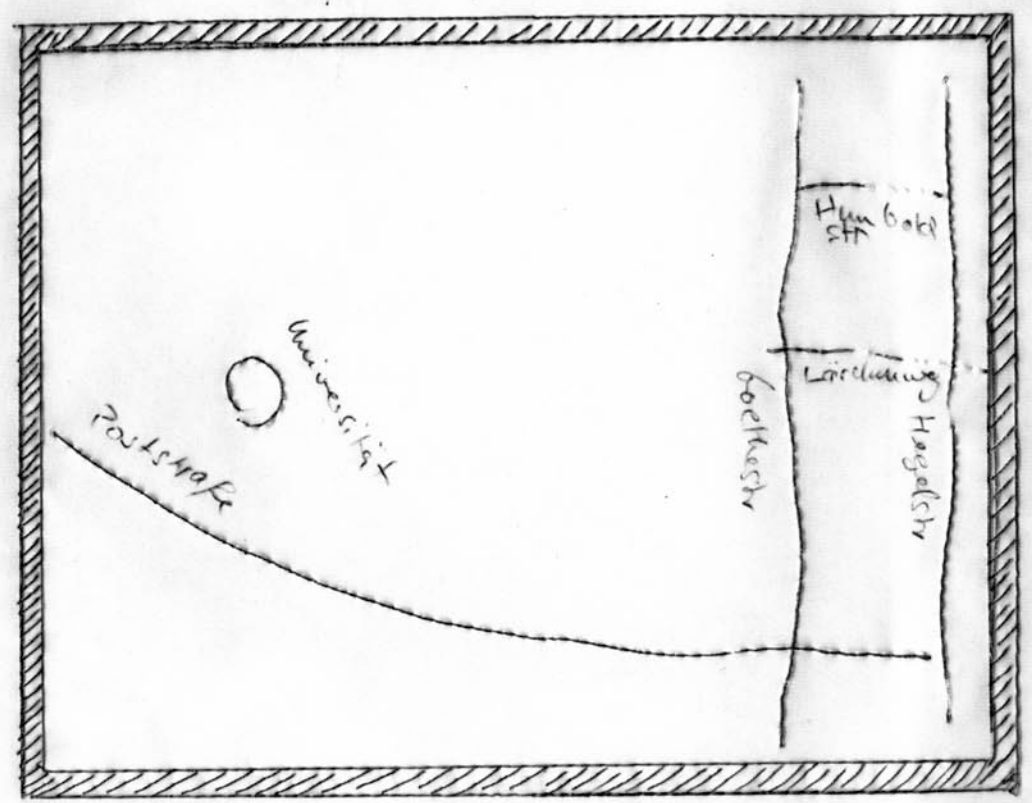
## Experiment 2: Wiederholung mit sehbehinderten Menschen

- Wiederholung von Experiment 1 mit blinden und sehbehinderten Menschen
- Der gleiche Ablauf (mit variierten Tests)
- 7 Teilnehmer ausgewertet (2 weiblich), 9 erhoben
- Altersdurchschnitt: 41.2 Jahre (*SD* 9.76)
- Geburtsblind (3), früh erblindet (2), spät erblindet (2)

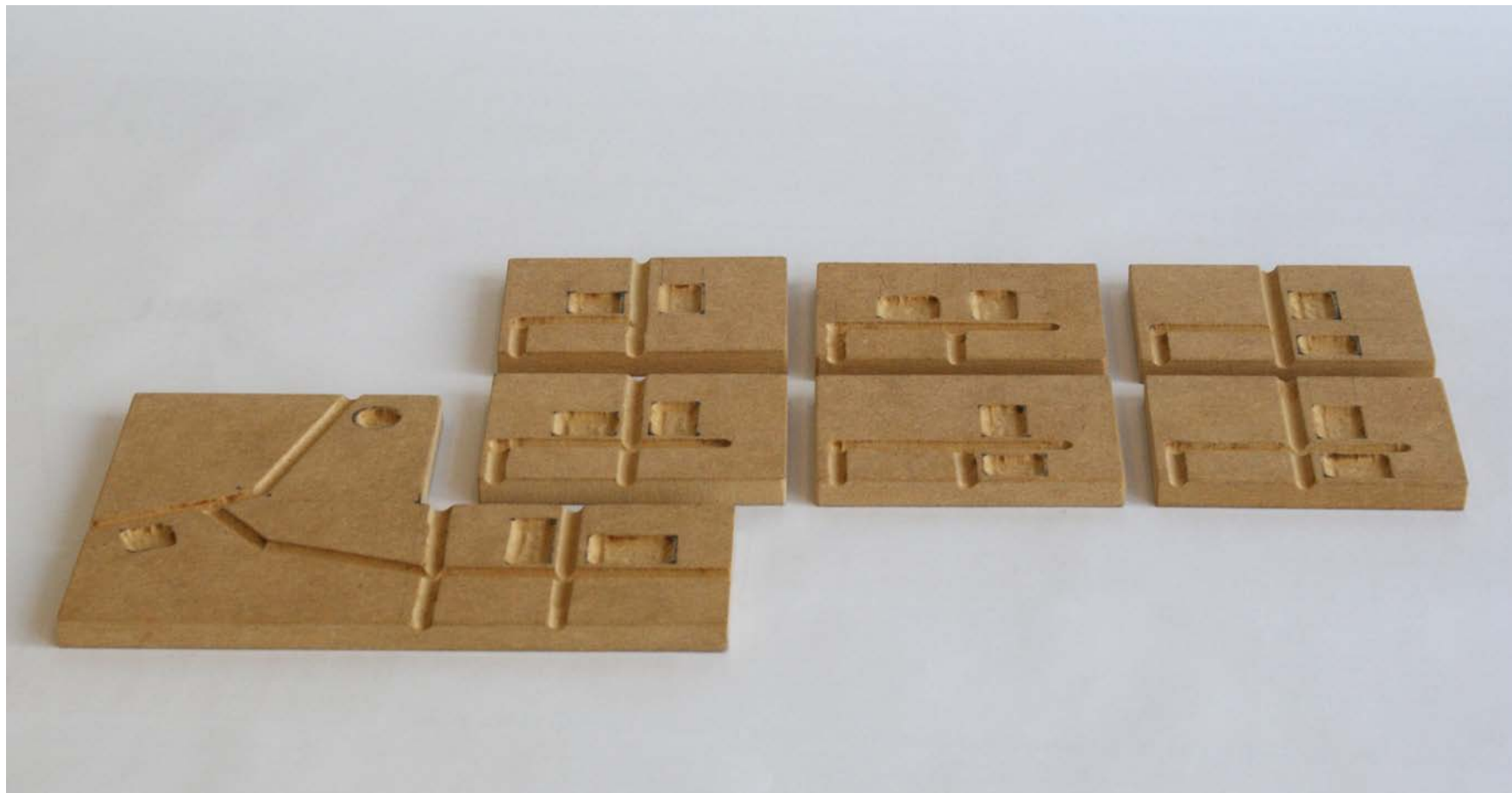
# Taktiler Zeichenbrett



# Beispiel: Sketch Map



# Taktiler Puzzle



Diese Folie kann online leider nicht zur Verfügung gestellt werden



Diese Folie kann online leider nicht zur Verfügung gestellt werden

Diese Folie kann online leider nicht zur Verfügung gestellt werden

Diese Folie kann online leider nicht zur Verfügung gestellt werden

Diese Folie kann online leider nicht zur Verfügung gestellt werden

Diese Folie kann online leider nicht zur Verfügung gestellt werden



# Was bedeutet das für ein Interface?

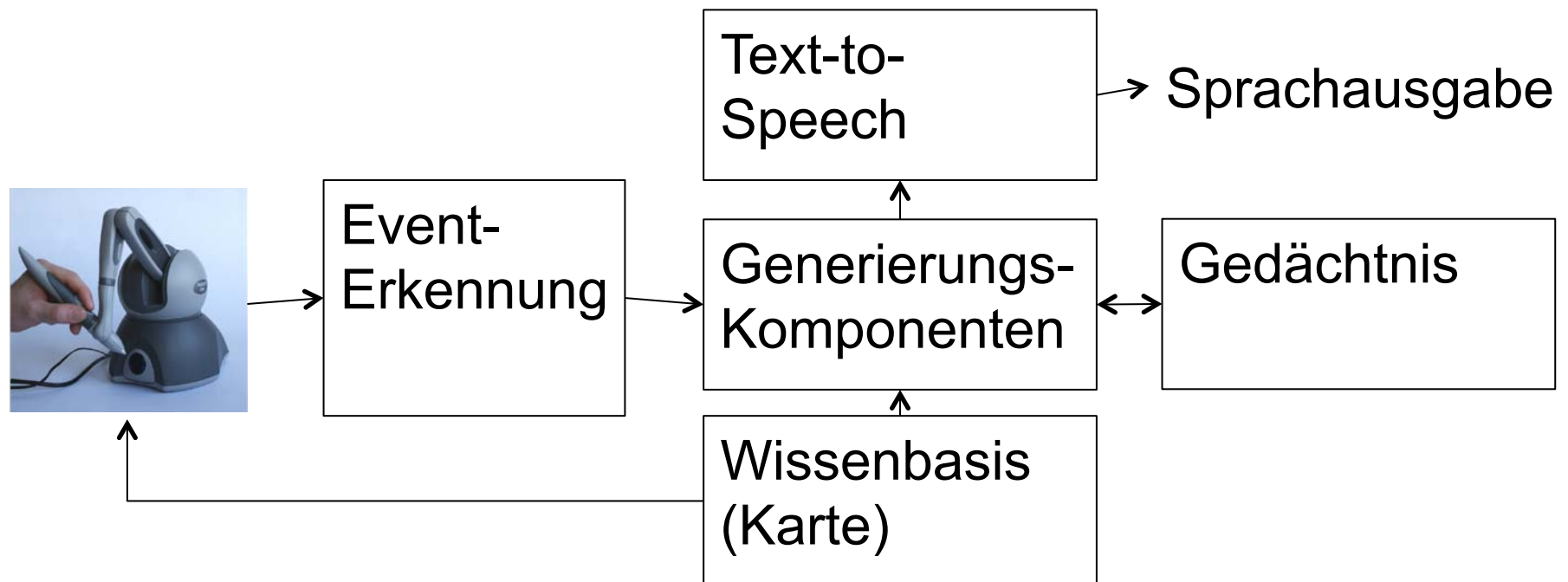
→ Verbal unterstütztes Interface ist signifikant effizienter

# Gliederung

1. Motivation & Technik
2. Was gesagt werden soll
3. Experiment 1: Effekt verbaler Assistenz bei Sehenden
4. Experiment 2: Effekt verbaler Assistenz bei blinden und sehbehinderten Menschen
5. VAVETaM-Prototyp
6. Ausblick und Zusammenfassung

# VAVETaM-Implementation

- Durch Simulation beweisen, dass automatisiertes System möglich ist (Grundlagen vorgeschlagen in Lohmann, Eschenbach und Habel 2011; Lohmann, Eichhorn, und Baumann 2012)
- Gesamt-Prototyp entwickeln





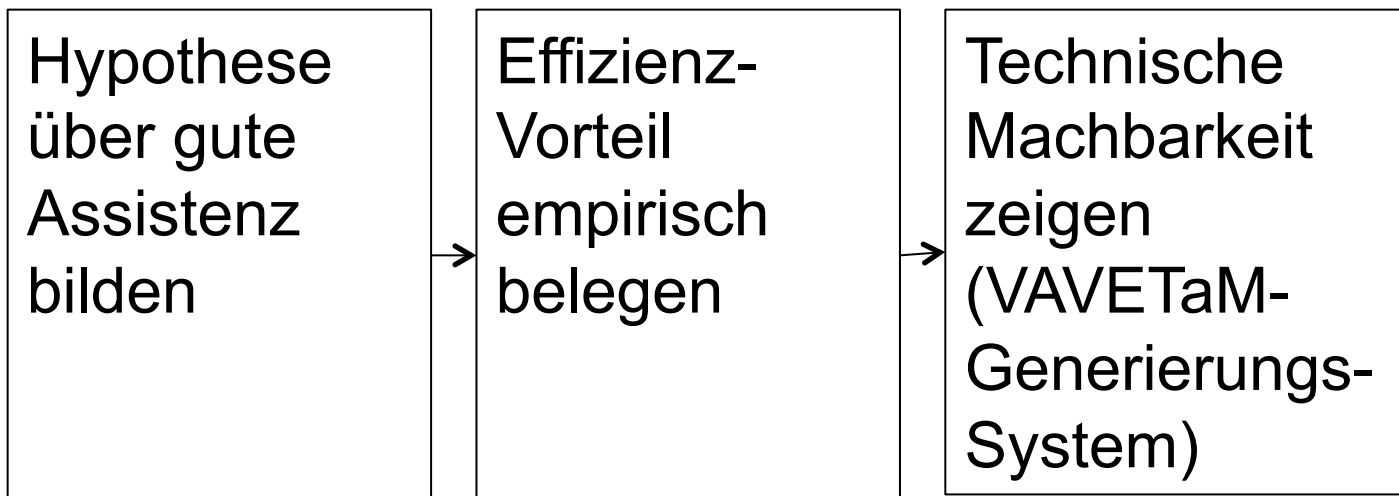
Video des laufenden Gesamtsystems.

?

# Gliederung

1. Motivation & Technik
2. Was gesagt werden soll
3. Experiment 1: Effekt verbaler Assistenz bei Sehenden
4. Experiment 2: Effekt verbaler Assistenz bei blinden und sehbehinderten Menschen
5. VAVETaM-Prototyp
6. Ausblick und Zusammenfassung

# Zusammenfassung



# Ausblick

- Benutzbarkeits-Studie mit VAVETaM-Prototyp
- Warum Unterschied zwischen Versuchspersonengruppen?
- VAVETaM-Implementation
  - Über Prototypen-Status hinaus verbessern
  - Verbessern der Dialog- und Satzplanung und des lexikalischen Zugriffs
  - Inkrementelle Sprachsynthese

**Danke für die Aufmerksamkeit!**

**Fragen?**

## **Acknowledgements**

- IRTG 1247 'CINACS' Cross-modal Interaction in Natural and Artificial Cognitive Systems

- Bortz, J., & Schuster, C. (2010). Faktorenanalyse. *Statistik für Human-und Sozialwissenschaftler*, 385–433.
- Giudice, N. A. (2004). *Navigating Novel Environments: A Comparison of Verbal and Visual Learning*. University of Minnesota.
- Giudice, N. A., Bakdash, J. Z., & Legge, G. E. (2007). Wayfinding with words: spatial learning and navigation using dynamically updated verbal descriptions. *Psychological Research*, 71(3), 347–358.
- Jacobson, R. D. (2002). Representing spatial information through multimodal interfaces. *Sixth International Conference on Information Visualisation (IV'02), London*.
- Lohmann, K. (2011). The Use of Sketch Maps as Measures for Spatial Knowledge. In J. Wang, K. Brölemann, M. Chipofya, A. Schwering, & J. O. Wallgrün (Eds.), *Understanding and Processing Sketch Maps – Proceedings of the COSIT 2011 Workshop*, ifgi Prints (Vol. 42, pp. 45–54). IOS Press.
- Lohmann, K., Eichhorn, O., & Baumann, T. (2012). Generating Situated Assisting Utterances to Facilitate Tactile-Map Exploration: A Prototype System. *Proceedings of the NAACL Workshop on Speech and Language Processing for Assistive Technologies 2012*. Presented at the NAACL Workshop on Speech and Language Processing for Assistive Technologies 2012, Montreal (QC).
- Lohmann, K., Eschenbach, C., & Habel, C. (2011). Linking spatial haptic perception to linguistic representations: assisting utterances for tactile-map explorations. In M. Egenhofer, N. Giudice, R. Moratz, & M. Worboys (Eds.), *Spatial Information Theory*, Lecture Notes in Computer Science (pp. 328–349). Berlin, Heidelberg: Springer.
- Lohmann, K., Kerzel, M., & Habel, C. (2010). Generating Verbal Assistance for Tactile-Map Explorations. In I. van der Sluis, K. Bergmann, C. van Hooijdonk, & M. Theune (Eds.), *Proceedings of the 3rd Workshop on Multimodal Output Generation 2010*. Dublin.
- Loomis, J. M., Klatzky, R. L., & Lederman, S. J. (1991). Similarity of tactual and visual picture recognition with limited field of view. *Perception*, 20(2), 167–177.
- Röder, B., Rösler, F., Grunwald, M., & Beyer, L. (2001). Ein Vergleich haptischer Wahrnehmungsleistungen zwischen blinden und sehenden Personen. *Der bewegte Sinn: Grundlagen und Anwendung der haptischen Wahrnehmung*. Berlin, Heidelberg: Springer.
- Wang, Z., Li, B., Hedgpeth, T., & Haven, T. (2009). Instant tactile-audio map: enabling access to digital maps for people with visual impairment. *Proceedings of the 11th international ACM SIGACCESS conference on Computers and Accessibility* (pp. 43–50). Pittsburg, PA: ACM.
- Wellek, S. (2003). *Testing statistical hypotheses of equivalence*. Boca Raton, London, New York, Washington, D.C.: CRC Press.