Representation of Faces and Places and the Influence of the Actor

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Wolfgang Röhrich
aus Ludwigsburg

Tübingen
2019

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	19.09.2019
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Hanspeter A. Mallot
2. Berichterstatter:	Prof. Dr. Bettina Rolke

Zusammenfassung

Sich in einer Umgebung zurechtzufinden, darin zu navigieren und mit anderen Personen zu interagieren ist ein alltägliches Verhalten. Beim Menschen liegen diesen Handlungen verschiedene Repräsentationen zugrunde - sowohl durch die Umgebung als auch durch die Identitäten der Personen, denen man begegnet. Es gab bereits umfangreiche Untersuchungen in diesen Forschungsbereichen, jedoch sind viele Ergebnisse dazu nicht eindeutig. Insbesondere wurde der Akteur selbst dabei oftmals vernachlässigt. Diesem sollte in dieser Doktorarbeit eine größere Beachtung zukommen. Darum wurde in den ersten drei Experimenten die Gesichtserkennung und -repräsentation untersucht. Im ersten Experiment wurden klassische Unterscheidungs- und Kategorisierungsaufgaben durchgeführt und mit den eigenen Gesichtern der Probanden erweitert. Es wurde erwartet, dass es an der Kategoriengrenze eine Erhöhung der Unterscheidungsleistung gibt. Im zweiten Experiment wurde der Raum, den verschiedene Gesichter aufspannen („face-space"), und darin natürlicherweise auftretende Verschiebungen untersucht. Es wurde angenommen, dass zwei Faktoren (die Ähnlichkeit zwischen den Gesichtern und die Markanz eines Gesichts) bei diesen Verschiebungen eine Rolle spielen. Im dritten Experiment wurde Priming der Identität getestet, sowohl mit Gesichtern der Probanden selbst als auch mit Gesichtern von anderen Personen, sowie der Einfluss von Namen.

Die Ergebnisse zeigten, dass es im ersten Experiment keine erhöhte Unterscheidungsleistung an der Kategoriengrenze gab, was durch den Grad der Vertrautheit mit den Gesichtern zu erklären ist. Allerdings konnte eine verbesserte Unterscheidung bei denjenigen Gesichtern festgestellt werden, die wiederum die Gesichter der Probanden selbst beinhalteten. Im zweiten Experiment wurden die natürlicherweise auftretenden Verschiebungen im Raum beschrieben, den die verschiedenen Gesichter aufspannen. Es gab eine Korrelation zwischen der Verschiebung und der Ähnlichkeit von Gesichtern bei Frauen. Die Ergebnisse des dritten Experiments zeigten künstlich induzierte Verschiebungen der wahrgenommenen Identität, welche im Einklang mit den verwendeten Primingreizen waren. Namen können ebenfalls Primingeffekte hervorrufen, allerdings unterscheiden sich diese Effekte wenn die Namen in der Anweisung verwendet wurden. Es wurde ein Primingmodell entworfen, das die zuvor erwähnten Ergebnisse berücksichtigt und die Reaktionszeitverläufe erklären kann. Dabei wurde auch ein Einfluss der Aufgabenanweisung erkennbar. Die Ergebnisse passen zum IAC-Modell der Gesichtserkennung und werden unter Berücksichtigung der Gehirnregionen, die bei der Gesichtserkennung beteiligt sind, diskutiert.

Im zweiten Teil dieser Doktorarbeit wurden zwei Experimente zur Untersuchung der räumlichen Navigation durchgeführt. Im vierten Experiment wurden Passanten an verschiedenen Orten in der Stadt befragt und darum gebeten einen von zwei bekannten Plätzen zu skizzieren. Dazu mussten sie die räumliche Anordnung und den Aufbau des Platzes aus ihrem Langzeitgedächtnis abrufen. Im fünften Experiment wurden weitere Passanten in der Stadt befragt und ihnen wurden Ansichten (Fotografien) von verschiedenen Plätzen der Stadt gezeigt. Sie sollten beurteilen ob es sich bei den jeweils gezeigten Bildern um einen ganz bestimmten Platz (Marktplatz) handelt oder nicht.

Die Ergebnisse zeigten, dass die Skizzen aus dem vierten Experiment zum einen eine Vorzugsrichtung aufwiesen - verursacht durch den Grundriss der Stadt - und zum anderen vom Befragungsort der Probanden hin zum Zielplatz ausgerichtet waren. Es war auch eine Regionalisierung erkennbar, welche sowohl durch den Grundriss der Stadt als auch durch die

Entfernung bedingt war. Die räumliche Repräsentation der Probanden an entfernten Befragungsorten unterschied sich von derjenigen in der Nähe des Zielplatzes. Experiment 5 zeigte, dass die Reaktionszeiten der Probanden abhängig von ihrer Entfernung zum Zielplatz waren und mit größerem Abstand kürzere Reaktionszeiten auftraten. Weiterhin wurden diejenigen Ansichten, die man zuerst sehen würde wenn man vom Befragungsort zum Zielplatz hin laufen würde, tendenziell schneller erkannt als andere Ansichten des Zielplatzes. Die Ergebnisse beider Experimente sind im Einklang mit dem „Viewgraph-Modell“, das hier ebenfalls diskutiert wird. In diesem Modell werden die Ansichten verschiedener Orte gespeichert und mit Handlungsanweisungen versehen, wie man von der einen Ansicht zur nächsten kommt.

Die Unterschiede und Gemeinsamkeiten von Gesichts- und Ortsrepräsentationen werden außerdem mit den involvierten Hirnregionen diskutiert. Wenngleich meistens unterschiedliche Regionen beteiligt sind gibt es dennoch Strukturen, die in beiden Fällen verwendet werden. Im Gegensatz dazu gibt es auch unterschiedliche Areale im Gehirn, die wiederum vergleichbare Funktionen in diesen beiden Aufgabenfeldern erfüllen. Schlussendlich kommt sowohl bei Gesichts- als auch bei Ortsrepräsentationen dem „Selbst“ eine wichtige Funktion zu. Es wird sowohl funktional als auch anatomisch separat behandelt und hat einen Einfluss auf andere Repräsentationen.

Summary

Navigating in an environment and interacting with other people is an everyday behavior. In humans, these actions have underlying representations both of the environment and of the persons' identities one meets. While there was already substantial research in these fields, many results are still inconclusive. Especially the actor oneself was often neglected. Hence this doctoral thesis is focused on the role of the actor in both face recognition and place recognition.

Face recognition and representation was investigated in the first three experiments. In experiment 1, classic discrimination and categorization tasks were conducted and extended with own faces of the participants themselves. An increase of discrimination performance at the category boundary was expected. In experiment 2, the face-space and natural occurring shifts within were investigated. Two factors (similarity between faces and distinctiveness of a face) were hypothesized to play a role in this shifting. In experiment 3, identity priming was tested with faces of the participants themselves and other faces as well as the influence of names.

Results showed that in experiment 1 there was no increased discrimination performance at the category boundary, which could be explained by the level of familiarity with the faces. However, an increased performance could be found for faces containing oneself. In experiment 2, natural shifts within the face-space consisting of the presented faces were described. A correlation between similarity ratings and shifts was present for female faces. Results of experiment 3 showed an artificially induced shift congruent with the priming stimuli. Names can also cause priming effects; however, these effects are different when names are used as instruction. A priming model was designed that considers the results of the aforementioned experiments and that can explain the response time curves, indicating an influence of the task instructions. The results fit to the IAC-model of face recognition and are discussed along with brain regions involved in face recognition.

In the second part of this thesis, two experiments were carried out to investigate spatial navigation. In experiment 4, passers-by were asked at different locations around town to sketch one of two well-known city squares. Therefore, they had to recall the squares' layouts from long-term memory. In experiment 5, passers-by around town were presented with views (photographs) of different city squares and they should judge if it was a certain one (market place).

Results showed that the sketches of experiment 4 had a general orientation bias - caused by the city layout - and were also oriented in relation to the participants' location towards the square. Results indicated a regionalization both by distance and the city layout. Participants' spatial representations at distant locations were different from those nearby the square. Experiment 5 showed that participants' response times differed in dependence on their distance from the square, revealing a negative correlation with distance. Furthermore, it was indicated that views that one would see upon approaching the square from the respective interview location were identified faster. The results of both experiments are supported by and discussed along with a viewgraph model. In this model the views of different places are stored and connected by instructions how to get from one view to another.

In addition, differences and commonalities between face and place representation are discussed along with the involved brain regions. While mostly separate, there are also structures that are recruited in both cases. In turn, there are separate areas in the brain that fulfill comparable roles

in these tasks. Finally, in both representations the "self" plays an important role. It is treated separately both functionally and anatomically and influences other representations.

Acknowledgements

First and foremost, I would like to thank Prof. Dr. Hanspeter A. Mallot for supporting me and my work at his chair of Cognitive Neuroscience through all the years. I'm thankful for all his ideas and suggestions for improvements.

I am grateful to Prof. Dr. Bettina Rolke for reviewing my work and supporting me with helpful tips and ideas.

I would like to thank PD Dr. Gregor Hardieß for his advice and suggestions for the experiments.

My thanks go to Dr. Marc Halfmann for programming the data collection on the tablet as well as the students Alisa Volkert, Esther Kutter, Niklas Binder, Julia Mayer, Frank Riemer and Amelie Haugg for their help with data collection for the different experiments.

This doctoral thesis was supported in part by the Werner-Reichardt-Centre for Integrative Neuroscience (CIN) Tübingen in the project "Being somewhere - spatial contents of self-representations and their role in spatial cognition".

Furthermore, I would like to thank all the colleagues at the Cognitive Neuroscience Lab for interesting discussions, helping with experiments and entertaining moments.

Special thanks go to Andrea for supporting me with ideas, testing the experiments, proofreading the thesis, thoughtful discussions and being there in general.

My deep gratitude goes to my brother and my parents. They always supported me and made this possible.

Contents

List of figures

1. General introduction

The recognition of the environment that surrounds us is a fundamental part in everyday life. This recognition happens through a multitude of senses and impressions. The locations at which interactions with the environment take place are often known to us. This means that there has to be an internal storage or (long-term) memory that holds information about said locations so that it can be compared to the sensory input and eventually identified and recognized as familiar. For humans, but also many animal species, vision is the most important sense for orientation and navigation, although, also other senses such as smells or sounds contribute to the identification of locations. Of course, there are other species that (almost) completely rely on those senses (e.g. cave animals) or even use senses (mostly) unavailable to humans (e.g. polarized light or magnetic field of the earth).
A rather simple strategy to memorize and recover a place is the so-called snapshot strategy, which has been shown to be used e.g. by honey bees (Cartwright & Collett, 1982), but can also be used by humans (Jacobs et al., 1998; Gillner et al., 2008; Lancier, 2016). In this snapshot strategy, the spatial position and coverage of landmarks on the retinal image are memorized. In retrieval, the actor moves until the current image and the stored image match again.
However, humans are also able to imagine familiar places and individual details within these places, too, which requires some kind of mental representation of these locations. Such imagery can be used to plan routes or direct other persons to some place. For an overview of navigation strategies see Mallot (2000).

1.1. Self-awareness

To be able to navigate in a (familiar) environment, not only knowledge of this environment but also knowing where "oneself" is located at the moment is essential. Furthermore, for complex navigation and actions one also needs to know where other objects and actors are located. At least for imagery, this requires a differentiation between oneself and others as there are no physical entities involved. This raises questions how "self" is defined, what it consists of and how it gets recognized. However, these questions accompany humanity for millennia, and already the Greek philosopher Socrates asked in Plato's work *Philebus*: "Shall we not say that our body has a soul?", and further: "Where did it get it?" It has been picked up by many other philosophers throughout the centuries and lately also psychologists and neuroscientists investigate the topic. Still, there is no holistic answer. However, partial areas that contribute to "self" could be identified and described.

A sub-part of "self" is self-awareness and the ability to recognize oneself. A well-established test for visual (self) recognition is the so-called "mirror test", or "rouge test" for humans, invented by Gallup Jr. for chimpanzees (Gallup Jr., 1970). In this test, a patch is attached to the human's or animal's face (or another body part that normally cannot be seen), and it is investigated if the individual starts to groom himself/itself and tries to remove that patch, which would mean he/it recognized the mirror image as oneself, or if the image is seen as another individual and the person/animal tries to remove the patch "from the mirror", or if other behaviors are carried out, e.g. interacting with a stranger. It has been shown that in humans this self-recognition starts to emerge at the age of about 20 months (Amsterdam, 1972). Several animal species have passed this test, too, e.g. chimpanzees (Gallup Jr., 1970), bonobos (Walraven et al., 1994), orangutans (Suárez & Gallup Jr., 1981) and gorillas (Posada & Colell, 2006; Allen & Schwartz, 2008), but also dolphins (Marten & Psarakos, 1994) and orcas (Delfour & Marten, 2001) as well as Asian elephants (Plotnik et al., 2006) and the magpie (Prior et al., 2008). There have been reports for African elephants passing the test (Birmelin), too. For reports of self-recognition in ants once they see themselves in a mirror see Cammaerts & Cammaerts (2015).

There are also other cues that allow for (self-) identification. For some animals, smelling plays an important role - sometimes more than vision does. Dogs typically do not pass the mirror test, however, it has been found that they are able to discriminate themselves from others by means of smell (Bekoff, 2001; Horowitz, 2017). Kozlowski & Cutting found in 1977 that humans can identify the sex of a person just by the walking pattern. For that they showed moving dots that were put on the joints of someone's body (point light walker). This does not allow for a full identification of a person unless he has a very specific gait (e.g. because of an accident) but can be used to narrow down the possibilities.
Also, inference from memory can be used for recognition: Typical habits like wearing a certain piece of clothing or certain gestures or other things are used for (self-) recognition, too. Finally, someone's name is a strong identifier for a person. Such other sources of information are also often used by patients with prosopagnosia who cannot rely on their face recognition (see below).
The more identity cues are present, the better the recognition works (Platek et al., 2004). This points to a self-recognition system that integrates information.

In this doctoral thesis, among others, the recognition of oneself and other persons was investigated by means of pictures of faces and names, as these provide typical and rich information about a person's identity. Human face recognition will be addressed further below. A second topic of this thesis was the effect of priming both in face recognition and spatial cognition.

1.2. Priming

"Priming" here means that a stimulus is processed differently if it follows on another stimulus which affects that one. Priming used in the current form is an implicit memory effect (Graf & Schacter, 1985; Roedinger, 1990; see Schacter (1987) for a historic review) and was first discovered and investigated in linguistics. It has been found that participants responded faster to words belonging to the same category if they were associated and semantically closer to each other (e.g. Schaeffer & Wallace, 1969; Meyer & Schvaneveldt, 1971) because of the previous experience with a word from that same category. The priming effect of words was also investigated in experiment 3 of this thesis. It has further been found that such priming effects are not only present for words but also for pictures, gestures, sounds, smells and other stimuli. Priming can be discriminated into discrete groups: Repetition priming, semantic priming, associative priming, response priming and affective priming, with potential new groups emerging as research goes on.

In *repetition priming*, the stimulus in question or similar stimuli are presented repeatedly. Eventually, the response is faster and/or more accurate, depending on the task.
In *semantic priming*, there is a semantic connection between the prime and target. For words an example would be "table" and "chair" as they belong into the same group "furniture"; see Neely (1991) and Lucas (2000) for reviews. If the first word is presented or thought of, similar words are

activated in the brain. If one of these is presented afterwards, participants respond faster compared to other words that are not connected.
Associative priming is related to semantic priming, however, the items do not have to be semantically related but only associated, which means it is likely that they occur together. An example would be “web” and “spider” for words.
Response priming typically investigates visuomotor processing with prime-target combinations with a short interstimulus interval and stimulus-onset asynchrony. Interstimulus interval describes the time between two stimuli, while stimulus-onset asynchrony describes the time between the beginning of stimulus one and the beginning of stimulus two, which means it includes the duration of the first stimulus as well as the interstimulus interval. Participants are requested to decide and respond as fast as possible. Depending on the addressed senses (visual, auditory, tactile, …), complexity of stimuli, timings and spatial locations of prime, mask and target stimuli, complexity of task and complexity of responses, and presence or absence of attention, different response patterns can occur. It is possible to cause facilitation with a certain prime and target stimulus, which means that the participants respond faster and/or more correctly. However, with the exact same prime and target stimulus combination one can also cause an inhibition (slower or false response) by only changing the stimulus-onset asynchrony between them. This allows for testing of conscious and unconscious brain mechanisms. This kind of priming was also used in several experiments of this thesis for face recognition and place recognition.
Affective priming uses existing or newly evoked emotions and moods to influence processing of information. Its effect, among others, depends on presentation duration of the prime and the presence or absence of access of consciousness (Zajonc, 1980; Murphy & Zajonc, 1993; Forgas, 1995).

With masking, a perceptual mask is applied before, after or in parallel to the prime stimulus. For words such masks often consist of other letters or nonsense words. The idea of masks is to limit the exposure of the stimulus to the exactly desired duration and cancel out afterimages or other stored memory information, or to investigate how such masks alter the processing in the brain and therefore the participants’ responses.
Sometimes the priming effect can be reduced or even vanishes if the participants are aware of the prime, e.g. if they think they might get influenced and therefore consciously respond in a different way as they would have otherwise. To avoid this, one has to limit the exposure to the prime stimulus either by reducing presentation time or by presenting the next stimulus - be it mask or target.
In certain paradigms, typically with short stimulus-onset asynchrony, the target stimulus itself can be used to mask the prime stimulus (as it has been done in experiment 3 of this thesis, too). Still, the priming effect is dependent on the stimulus-onset asynchrony (Vorberg et al., 2003). This strongly depends on the nature of the stimuli.

1.3. Scientific issue

The representation of the environment that surrounds us along with its entities that can be interacted with is a fundamental part of everyday life. Many studies have investigated separate segments of these representations, yet, there are varying results and conclusions.

The first topic of this thesis was focused on face recognition and representation. Previous studies (see below) had ambiguous results regarding the correlation between discrimination and categorization performance, and therefore categorical perception. For discrimination, two slightly different stimuli are presented followed by a third one that is either the same as the first or second one. For categorization, participants have to decide when a stimulus switches from one to the other category. It is said that at the categorical boundary of two items, the discrimination rate is best since the representation switches from one to the other. This, of course, requires that representations of both items exist. Therefore, experiment 1 was conducted with classic discrimination and categorization tasks like they were carried out by Beale & Keil (1995), who found an increased discrimination performance at the inflection point of the curve of the categorization performance - however, only in famous (= familiar) faces. The inflection point occurs where the participants' answers switch from one to the other item, which means at the categorical boundary (see e.g. Figure 8 for two curves with inflection points at the 50 % level). Other studies, in turn, found an increase also for unfamiliar faces (e.g. Levin & Beale, 2000), or did not find an increase for sex judgements at all (Bülthoff & Newell, 2004). Additionally, the first experiment presented in this thesis was extended to also investigate the role of the participants "themselves" in these tasks, and how familiarity levels influence the response behavior. This was accomplished by adding their "own" face as stimulus and contrasting it with pictures of "other" faces. It is hypothesized that one's own identity or "self" plays a special role in face representation, resulting in a different performance in these tasks.
Faces differ from each other in certain features (e.g. nose, chin, eyes, …). Faces and their features thus can be described through a multidimensional face-space, which they span with their different features. In experiment 2, this face-space and naturally occurring shifts in said space were investigated along with two factors (distinctiveness of and similarity between faces) that were hypothesized to influence these shifts. This was also in preparation for experiment 3 in which artificial shifts were induced.
In experiment 3, face recognition of "self" and "other", as mentioned above, was combined with a priming setup. Again, on the one hand to replicate previous studies that had inconclusive results, like e.g. Bruce & Valentine (1986) and Ellis et al. (1987) who found no priming effects of names, while Calder & Young (1996) did find priming effects, and on the other hand to extend it to further investigate whether "self" has a special role in face representation by introducing pictures of "self" and "other" faces.

The replication of previous studies and analysis of the results with this setup was the first aim of this thesis. It was extended to investigate the role of "self" in these tasks. A priming model that accounts for these results was developed and is discussed along with shifts in face-space and with current models of face recognition and representation, which was the second aim of this thesis. Face-space is explained and natural occurring shifts within this face-space were investigated in experiment 2. In experiment 3, these shifts were artificially induced by priming.

In the second part of the thesis, and as a third aim, the representation of places was investigated and how this representation is influenced by "self" - here one's current location. Two experiments were accomplished in which the first one (experiment 4) examined differences in recalling nearby and distant locations from long-term memory and how participants' proximity affected the representation. A view-based model, suggested by Schölkopf & Mallot (1995), is presented and explained how familiar regions could be stored in long-term memory, considering the results of the experiment.

The second experiment (experiment 5) was focused on the nature of the aforementioned "views" and whether visual external cues could yield a certain representation of a location, similar to how imagined views in experiment 4 and in a study by Basten et al. (2012) were able to do so.

The results of both experiments and the view-based model are discussed with other models of spatial navigation.

Priming was used in the experiments of this thesis in several ways. First, in face recognition, both pictures of faces and names were used to trigger priming effects. Second, in spatial navigation, imagination of a certain place or showing pictures of it was used to prime mental images of that location (and associated features of it).

In the following, three experiments regarding face recognition will be presented and different aspects will be discussed along with models of (primed) recognition. In the second part of this thesis, spatial cognition will be investigated. Two experiments will be presented and their results will be discussed and compared to a proposed model of wayfinding. Finally, differences and commonalities of the recognition and representation of faces and places will be presented.

2. Introduction for experiments 1 to 3

The human face is a unique composition of individual features. Even identical twins that hold the same genome and look confusingly similar still have little features they can recognize and from which they can tell each other apart. Furthermore, people are able to quickly find and recognize themselves on a picture showing a group of people, and they can identify themselves with "the face in the mirror". These insights demonstrate interesting aspects of both face recognition and self-identification with a face.

Human face recognition starts early in child development. Already newborns can mimic facial expressions (Field et al., 1982), although, not before the age of about 7 months they begin to understand the meaning of basic emotional expressions, e.g. a happy and a fierce face (Striano & Vaish, 2006; Leppänen et al., 2007; Hoehl & Striano, 2008). More sophisticated face processing like recognizing oneself in a mirror (see general introduction) starts at the age of 20 months (Amsterdam, 1972).

It is known that the recognition of another person by their face is impaired in patients with prosopagnosia. The phenomenon was first described by Bodamer in 1947 in three patients who lost their ability to recognize other people after brain damage had occurred. Later on, different types of prosopagnosia were described. In *apperceptive prosopagnosia*, the ability to recognize individual faces is impaired. Such patients may identify people by other sources of information like their voice or name or biographic knowledge but not by their visual face. In *associative prosopagnosia*, patients cannot tell if they know the person or report biographic information from the face, but they are able to tell the sex or age of the person from the face (De Renzi & di Pellegrino, 1998; Anaki, 2007). They are also able to make judgments between two or more faces if they are similar or different. Prosopagnosia can be acquired through head traumata but can also be already present in newborn, named *congenital prosopagnosia*. These findings indicated back then that the source of the disorder can be found in the brain, hence involved brain structures have been investigated in various studies.

The predominant anatomical structures for face recognition have been identified in the occipital and temporal lobes of the brain, especially the fusiform face area (see Kanwisher & Yovel (2006) for a review) in the lateral fusiform gyrus, and the occipital face area in the inferior occipital gyrus, which both, but not exclusively, have large impact on the detection of faces as well as identification (see general discussion of this thesis and Rossion et al. (2012) for an overview). They found that the fusiform face area responds to (low-level) facial features and is supposed to detect faces from other objects. This structure also gets activated early in visual processing in cases of pareidolia (Hadjikhani et. al., 2009), which means that familiar patterns are "seen" although they are not there. This happens often with "faces" in non-face objects, e.g. a yellow smiley (which is actually a circle with dots and lines) or certain rock formations (like the "Old Man of the Mountain" in New Hampshire, USA) or clouds or the "Man in the Moon" etc. However, pareidolia is not restricted to faces and also works with animals or other structures as can be seen with the "Elephant Rock" on Heimaey island in Iceland. Rossion et al. (2012) also stated that the fusiform face area is not very face-selective as it was activated by cars and scrambled cars, too. Still, the largest activation was present when normal faces were shown. The occipital face area, in turn, showed a higher degree of face-selectivity in their experiments. Pitcher et al. (2011) stated that the occipital face area is among the first stages of processing of face perception and "preferentially represents the parts of a face, including the eyes, nose, and mouth", however, not the spacing between them. Liu et al. (2010) were focusing in their study on the differences between occipital face area and fusiform face area. They concluded that the occipital face area is only sensitive to real face parts while the fusiform face area is sensitive both to faces parts and face configuration. This finding is supported by Zhang et al. (2012) who suggest that faces are represented in the fusiform face area as a "whole that emerges from the parts".

Other brain regions are involved in analyzing faces, too: The amygdala, for instance, plays an important role to detect emotions and trustworthiness - even if faces are seen only briefly and not perceived consciously (Freeman et al., 2014). It is also involved in recognition, learning and habituating to new faces, along with the hippocampus (Heit et al., 1988; Fried et al., 1997; Blackford et al., 2012). The "hippocampus supports the additional contextual element of where and when the item was encountered" (Bird & Burgess, 2008). Depending on the familiarity with the other face even more brain networks are involved (see e.g. Leveroni et al., 2000; Taylor et al., 2009). In the general discussion of this thesis this topic gets revisited.
Altogether, this allows for a rich representation of faces with additional knowledge of the persons' identities and associated information (see below). All these findings indicate that face recognition is not a simple single mechanism but rather a complex one. It has become subject of many studies and different models have been proposed how recognition of a person works.

Yin already found in 1969, when investigating upside-down faces, that there has to be "a special factor related only to faces" (Yin, 1969). Later, it has been found that faces are represented holistically (Tanaka & Farah, 1993; Farah et al., 1998; Wiese et al., 2013) and are discriminated from other objects at an early stage of perception (Taubert et al., 2011). However, this holistic processing is not rigid and modulated by other factors such as emotions (Curby et at., 2011). They found that faces get processed less holistically if negative emotion cues are present. Furthermore, it has been shown that faces are perceived categorically (see below and experiment 1), be it with respect to identity (Beale & Keil, 1995), expression (Calder et al., 1996 b), race (Levin & Angelone, 2002) or sex (Armann & Bülthoff, 2012). Though, categorical effects were not always found. The most debated factor here is familiarity. Already Beale & Keil (1995), but also Bülthoff & Newell (2004 and 2005), and Armann & Bülthoff (2012) found categorical perception for familiar but not for novel faces. Yet, Campanella et al. (2001 and 2003) also found categorical effects in recognition of unfamiliar faces; so did Levin & Beale (2000) in newly learned faces. To investigate these effects, the above mentioned studies created artificial continua of and between faces by using a morphing technique in which linear transitions of metrics and texture from one face to another were applied. In the following experiments, the same technique was used to create pairs of morph pictures of faces (see 'general material and methods for experiments 1 to 3' below).

A common model of face recognition named "interactive activation and competition (IAC) model of face recognition" (Figure 1) that considers the clinical and behavioral findings was proposed by Bruce & Young in 1986 and since then extended several times e.g. by Burton, Bruce & Johnston (1990), Burton & Bruce (1993) and Burton, Bruce & Hancock (1999). This model follows the basic ideas of an IAC model proposed by McClelland & Rumelhart (1981) for letter perception. In the face recognition case, the model proposes several independent sub-units which fulfill different tasks: Face recognition units (FRUs) are meant to visually recognize individual faces of persons. Subsequently, the according person identity node (PIN) gets activated, which carries information about "whom the face belongs to" as well as serve as a hub that connects to other units like name recognition units (NRUs), which get input from word recognition units (WRUs), and semantic information units (SIUs). Viewing a familiar face on a photograph would excite this person's FRU and then activate the person's PIN, which links to certain SIUs holding additional information about the person, like the person's name, occupation and hobbies. Finally, one is able to report this information via lexical output.

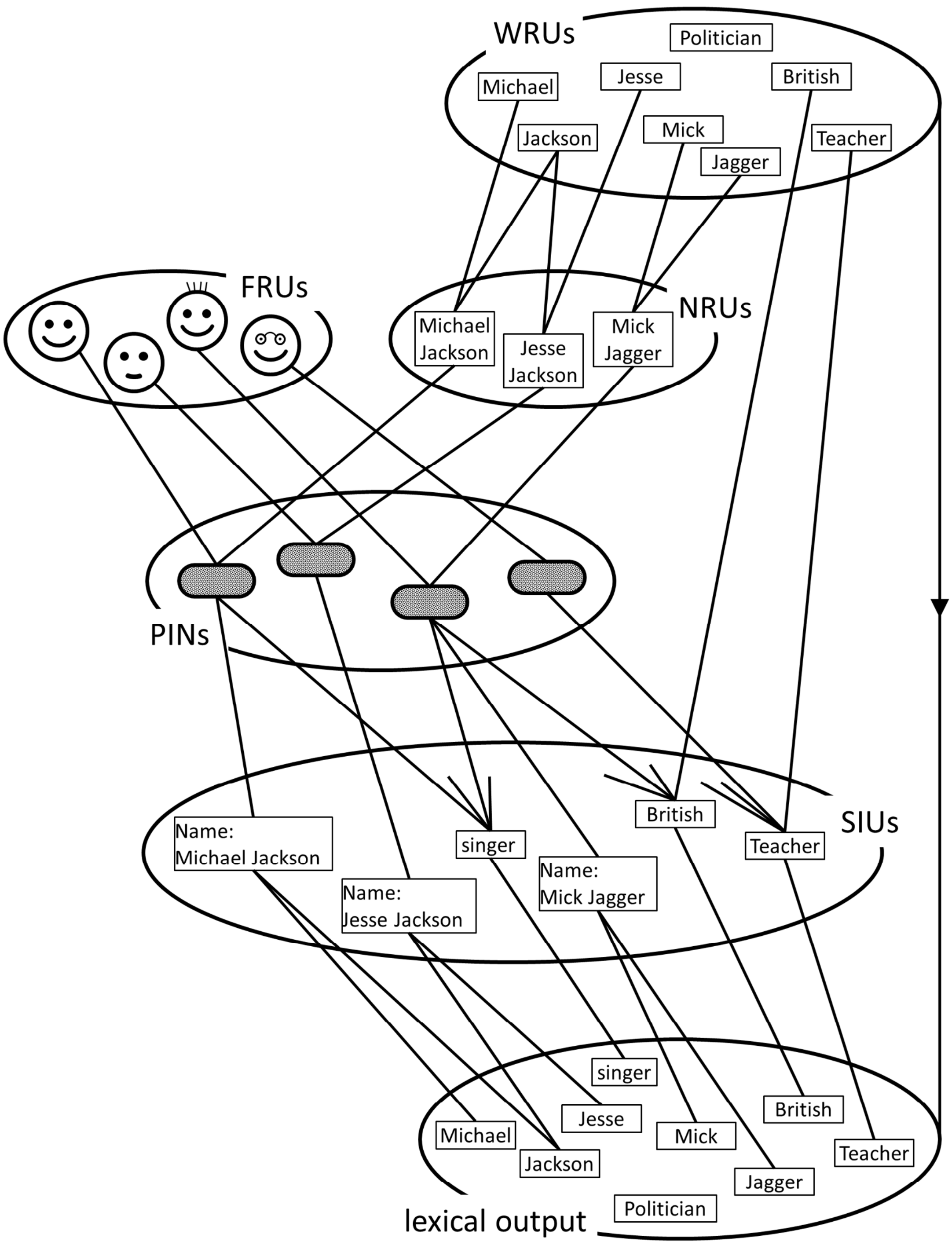

Figure 1. Extended IAC model by Burton et al. (1993). Different recognition units (FRU = face recognition unit, NRU = name recognition unit, WRU = word recognition unit,) forward information to the person identity nodes (PINs) in which the identity of a person gets determined. Known semantic information about that person is activated in the SIUs next, and finally, depending on the task, a lexical is output generated. (Adopted from Burton & Bruce, 1993).

The connections are bi-directional, which means activation can spread into both directions once one of it received a (matching) input. It has been shown that presenting a picture of a famous face, like the actor Oliver Hardy from “Laurel and Hardy”, would not only yield information about this actor but also on his counterpart Stan Laurel, leading to faster recognition of him on a picture or his name (Bruce & Valentine, 1986). This can be explained by the shared activated parts in the SIUs which link both persons together.
In the above stated example with the two actors (Laurel and Hardy), an example of “identity priming” was given. In fact, various studies have investigated the effects of priming on various parts of the recognition of faces, names and identities. In 1971, Meyer & Schvaneveldt discovered a facilitation effect in word recognition if pairs of words were associated with each other (e.g. doctor and nurse). Ellis et al. (1987) investigated repetition priming of face recognition in several experiments. They found that “recognizing a face as familiar is primed by prior exposure to a photograph of that person’s face but not by prior exposure of the person’s written name”. Also, the priming effect was present not only for famous persons from the media but also for those who were personally familiar to the participants, and the priming effect was strongest if the prime and test photographs were the same, and weakest if the two photographs were visually dissimilar (but still showing the same person).

Calder et al. (1996 a) and Calder & Young (1996) investigated “self-priming” with distinctive and caricatured faces as well as different prime forms (within and across domain, see below). Their findings included that caricatured faces, which hold more characteristics of a certain person, also lead to stronger priming effects. Furthermore, cross-domain (picture - name) priming was weaker than within-domain (name - name). They also found a weak priming effect for names on faces - in contrast to Ellis et al. (1987).

Both Dell’Acqua & Grainger (1999) and Eimer & Schlaghecken (2003) focused on unconscious and subliminal priming from pictures or pictorial cues with prime presentation times of no more than 17 ms. The priming effect strongly depended on congruency between prime and target (e.g. semantic category), stimulus onset asynchrony and presentation location (central or peripheral in field of view). Tong & Nakayama (1999) found that pictures of self faces were always identified faster in comparison to other faces in a visual search task. Pannese & Hirsch (2010) also investigated the self-face advantage in a priming experiment where subjects should judge sexes. They found priming effects in congruent stimulus presentations and a specific priming effect for self faces at short prime durations, too.

Although face recognition was investigated in all these studies, certain properties are still inconclusive. In the following three experiments, face recognition was investigated regarding face discrimination and categorization (experiment 1) as well as identity priming with pictures and names with focus on differences between “self” and “others” (experiment 3). Furthermore, shifts within the face-space between faces were examined (experiment 2).

3. General material and methods for experiments 1 to 3

In experiments 1 to 3, which investigated face recognition and representation, photographs of participants were presented on a computer monitor and printed cards. The acquisition and editing process for these pictures was the same for all of them, therefore, the method will be described in general hereinafter.

Biometric (frontal shot, neutral expression) high-quality color photographs of the participants' faces were taken with a digital camera (PowerShot G7, Canon, Japan) under controlled lighting conditions prior to the experiments. With Adobe Photoshop 7 (Adobe Systems, USA) individual masks were applied to all photographs to remove the background, the torso below the neck and the persons' head hair. The background was filled with black color then. Available hairstyle information results in different processing of a face in terms of identity and sex for both familiar and unfamiliar persons (Goshen-Gottstein & Ganel, 2000; Ganel & Goshen-Gottstein, 2002) and was therefore removed. Afterwards, a generic cranium was added to preserve a natural shape of the head. Images were cropped and resized to 500 x 500 pixels in which the head was covering about three quarters of the image (Figure 2).

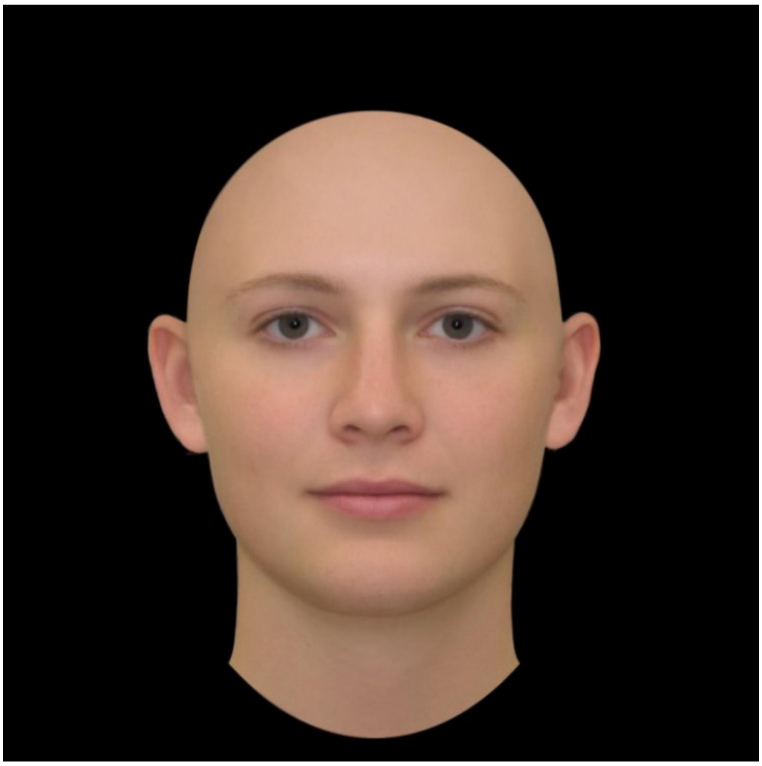

Figure 2. Example of a participant's face after processing as it was shown later in the experiments as stimulus.

3. General material and methods for experiments 1 to 3

Gender specific morph sequences with 10 % intervals were created with FotoMorph (Digital Photo Software, Norway) between all faces for experiments 1 and 3. For experiment 2, more intervals were selected, resulting in smaller steps of 4 % between images.

To create such morph-sequences, corresponding features in two to-be morphed faces were selected by hand (e.g. tip of nose, pupils, outline of the head) and a Delaunay triangulation was applied, leading to a mesh of triangles. Intermediate pictures were created by the underlying algorithm by shifting the previously selected feature points and areas in between and mixing pixel color according to the weighting of the two input pictures (e.g. an intermediate picture with 20 % of picture A and 80 % of picture B had feature points shifted for 80 % along the linear transition from A to B as well as 20 % color information from A and 80 % color information from B, resulting in a 20:80 % morph picture). Thus a smooth transition between two images was created (Figure 3).

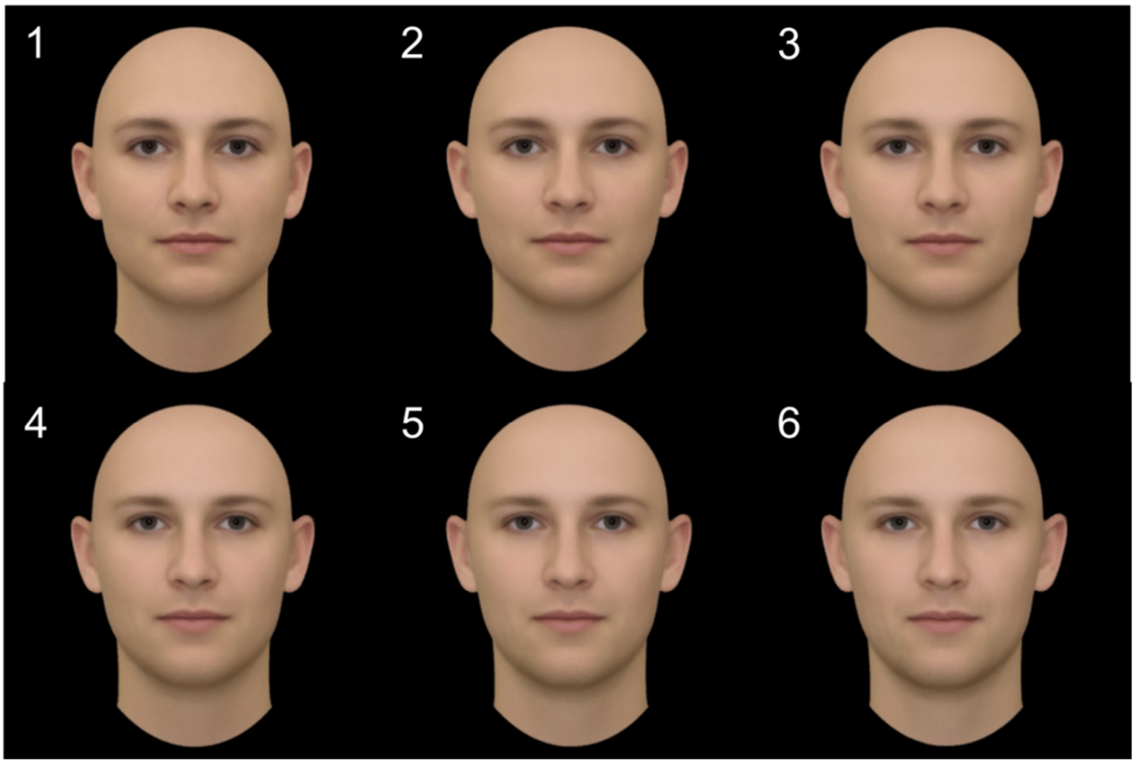

Figure 3. Example of a morph sequence. The first (1) and the last (6) picture are "pure" unmorphed images of two participants, pictures 2 - 5 are morphed intermediate pictures of these two, here with an interval of 20 % morph-level.

Faces and morphs were displayed on a PC monitor (1280 x 1024, 19", Iiyama B1902S) during the experiments. Prior to the experiments, the monitor delay was measured with an oscilloscope and a photodiode. The delay of the setup was constant with about 20 ms between trigger and display onset.

3.1. Analyses and statistics for experiments 1 to 5

All data from experiments 1 to 5 were analyzed with MATLAB R2012a (MathWorks, Natick, MA, USA) or IBM SPSS Statistics v22 (IBM Corporation, Armonk, NY, USA) and results plotted with MATLAB R2012a. All deviations in the text and figures (e.g. 3.5 ± 1.1 s) indicate standard deviations unless stated otherwise. All ANOVAs tested main effects and possible interactions. Significant differences between means were indicated with stars in the figures, where * = $p < 0.05$, ** = $p < 0.01$ and *** = $p < 0.001$.

3. General material and methods for experiments 1 to 3

4. Experiment 1 - Discrimination and categorization of faces

Categorical perception is a processing technique of the brain to narrow down the amount of information and focus on important differences by exaggerating and detecting boundaries between stimuli. This is not limited to visual impressions but applies to semantic differences in general, indicating a general cognitive process (Harnad, 1987), e.g. to discriminate between groups of objects, phonemes (Liberman et al., 1957 and 1967), colors (Bornstein & Korda, 1984), and other stimuli that can be grouped into discrete categories.

In this experiment, classic discrimination and categorization tasks were carried out, both with pictures of the participants themselves and with pictures of other persons. It was hypothesized that an increase of discrimination accuracy can be found at the category boundary and that it should be congruent with the inflection point of the categorization task - like Beale & Keil (1995) found in their study (see discussion). The main aspect of this experiment was the role of the participants' own faces and how they influence the discrimination and categorization performance. On the one hand, the participants' faces are well-learned and memorized faces. On the other hand, the familiarity of the other faces was initially low and familiarization happened through the course of the experiment, which can be sufficient to cause categorical perception effects (Levin & Beale, 2000). Therefore, it was further investigated whether these faces used in this thesis as stimuli can elicit categorical effects at the given familiarity level, too, and how morph sequences of mixed familiarity levels (pictures showing the participants themselves and other persons) influence the participants answering behavior.

4.1. Material and methods

4.1.1. Participants

Ten adults of European descent (five male, five female) aged between 22 and 30, with normal or corrected to normal vision participated in this study. A fixed amount of 8 € per hour was paid for participation. Care has been taken to select men without facial hair and women with no apparent facial make-up since morphing would have led to unrealistic pictures. Participants were informed about the collected data, the general procedure and that they were free to terminate their participation at any time. Written informed consent procedures adhere to the guidelines of the Declaration of Helsinki and was obtained from all participants.

4.1.2. Setup

Participants were placed in front of a PC screen (see general material and methods above) in a distance of about 50 cm in a dimly lit room. The room lighting was controlled and kept constant during the experiment. Stimuli were presented with Psychtoolbox 3 in MATLAB R2012a (MathWorks, Natick, MA, USA).

Images covered about 18 cm of the computer screen, which equals about 20° of visual angle at the given distance.

4.1.3. Procedure

The experiment consisted of two blocks: In the first block, each participant was asked to complete discrimination tasks. Afterwards, in the second block, categorization tasks had to be solved. In both blocks, participants were shown pictures of morph sequences of two kinds:

Condition 1: Three strangers morphed separately with a picture of the participant.
Condition 2: Three strangers (the same) morphed with each other.

4.1.3.1. Discrimination task

Each trial started with a fixation cross, which was displayed for 1 s. Then, participants were shown a sequence of three pictures. The first picture "A" was randomly taken from a morph sequence (see above) and presented on the screen for 0.75 ms. Next, the screen went black for 1 s (inter stimulus interval). Then, the second picture "B", taken from the same morph sequence but differing from picture "A" by 20 %, was presented for 0.75 s. Another interstimulus interval with black screen for 1 s occurred. Afterwards, a third picture was displayed for 1 s, and then the screen went black again. This third picture was one of the previously shown pictures "A" or "B". Participants should decide and indicate by pressing one of two buttons on the keyboard whether the third picture was the same as either the first one or the second one in this two-alternatives forced choice (2AFC) paradigm (Figure 4). After key press a new trial started. Each morph sequence between two persons consisted of eleven pictures differing by 10 % (see general material and methods for experiments 1 to 3). Since pictures "A" and "B" were selected and shown that differed by 20 % from each other, one obtained nine image pairs to describe the morph sequence from one end to the other. Each participant had to complete 216 trials (2 conditions x 3 morph sequences x 9 image pairs x 4 repetitions).

4.1.3.2. Categorization task

At the beginning of the categorization task, participants were shown onetime the three unmorphed faces he or she was morphed with until a key was pressed. Participants should look at and memorize these faces carefully for as long as they wanted. In the upper row the faces of the three other persons were shown; in the lower row the face of the participant him- / herself was shown (Figure 5). Between trials there was a black screen for 1 s as inter trial interval. The trial started with a fixation cross displayed for 1 s. Then, a picture randomly chosen from one of the six morph sequences (three from condition 1, three from condition 2) was shown for 1 s, followed by a black screen interstimulus interval for 1 s. Afterwards, two small pictures were displayed on the sides of the screen, showing the two unmorphed faces of the morph sequence the previous face was taken from. The participant then indicated by button press if the single face was more like the face on the left or right side ("Ist das eher die Person auf der linken oder der rechten Seite?") and a new trial began (Figure 5). Participants had to take a total of 198 decisions (2 conditions x 3 morph sequences x 11 pictures x 3 repetitions).

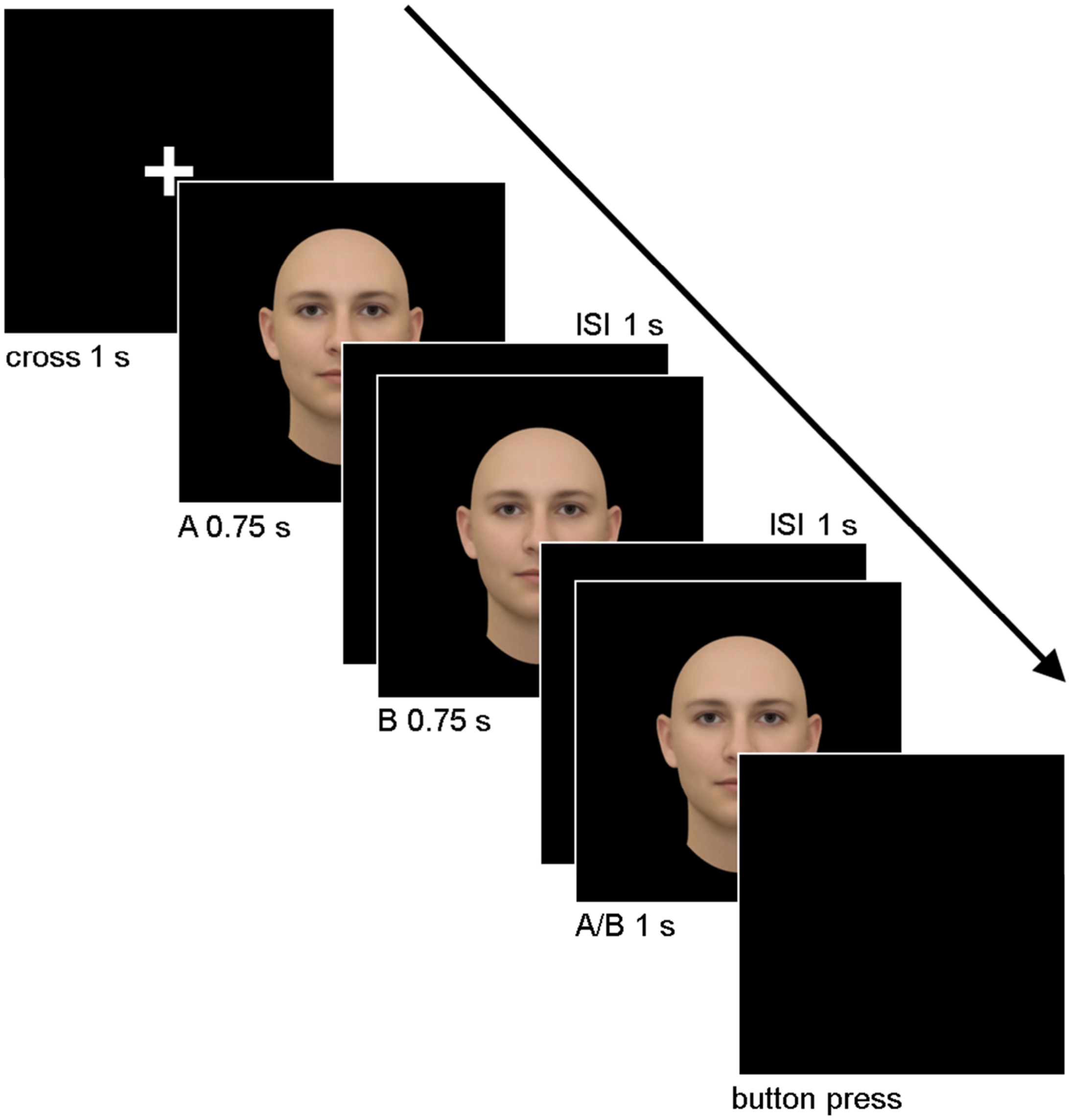

Figure 4. Sequence of a discrimination trial. Each trial started with a fixation cross for 1 s. Afterwards, the first picture (A) appeared for 0.75 s. Then, there was a black screen as inter stimulus interval for 1 s. The second picture (B) was displayed for 0.75 s. There was another interstimulus interval (ISI) of 1 s. Then, either the first (A) or second (B) picture was shown again for 1 s. After that the screen went black. The participant should indicate with a button press if the third picture was the same as either the first or the second one shown.

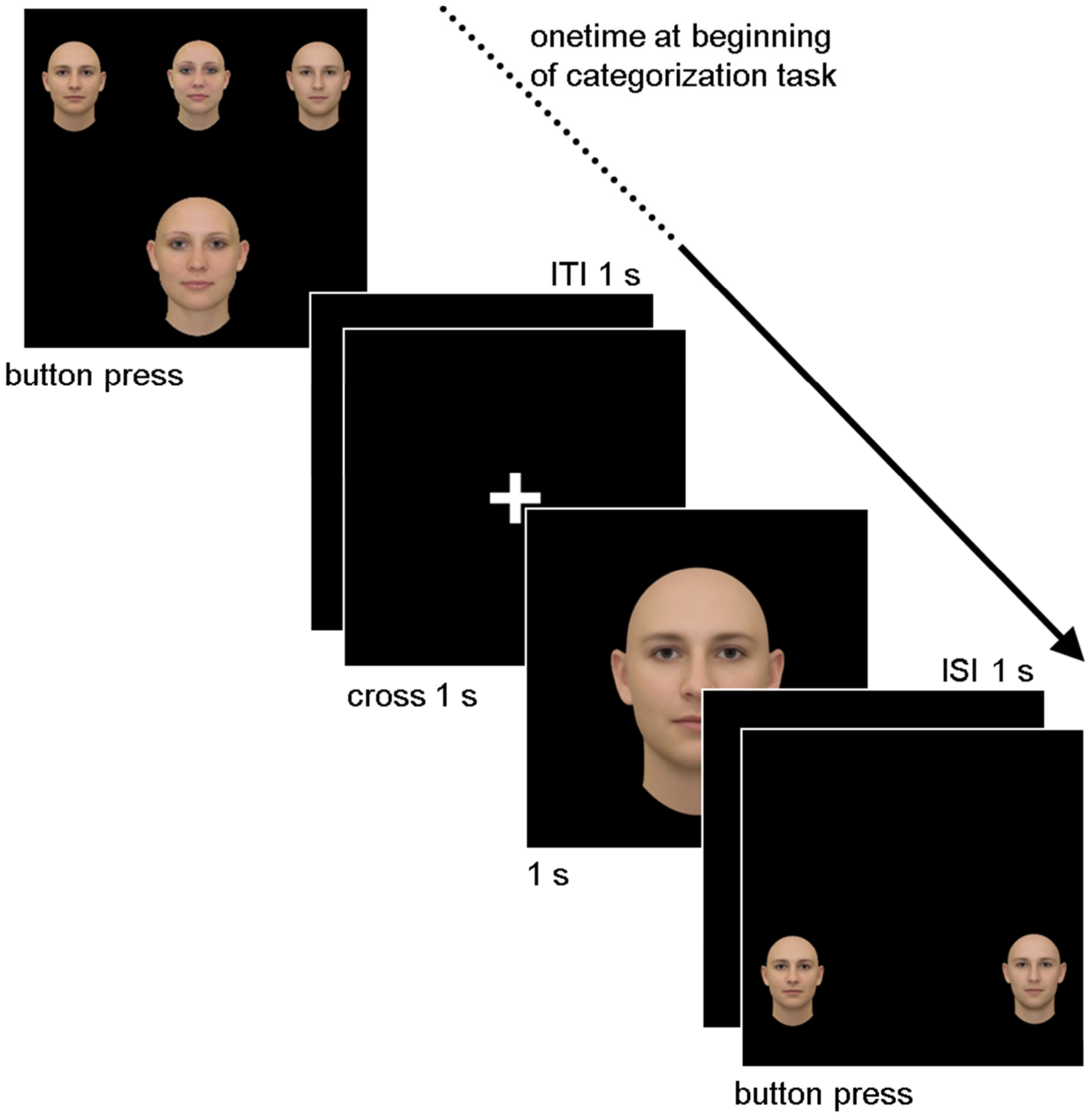

Figure 5. Sequence of a categorization trial. At the beginning of the categorization task, an overview with all faces was presented onetime for as long as the participants wanted. In the upper row the faces of the three other persons were shown; in the lower row the face of the participant him- / herself was shown. Participants were encouraged to carefully look at the faces and memorize them. Between each trial there was an inter trial interval (ITI) of 1 s. Each trial started with a fixation cross for 1 s. Afterwards, one picture of a morph sequence was shown for 1 s. Then, there was a black screen as inter stimulus interval (ISI) for 1 s. After that, in the lower left and right parts of the screen the endpoints (= unmorphed images) of the morph sequence were shown. The participants should indicate with a button press (left or right key) if the previously shown (morphed) picture was more like the person on the left or on the right side.

4.2. Results

Individual mean response times for both tasks (discrimination and categorization) were calculated for all participants. Responses that took longer than three times the standard deviation of the average were discarded. This happened in 42 out of 2160 trials in discrimination tasks and 44 times out of 1980 trials in categorization tasks.

4.2.1.1. Discrimination task

On average, the percentage of correct decisions was 71.92 ± 3.82 % for morph sequences that contained the participants' faces (condition 1) compared to 70.12 ± 2.69 % for sequences without the participants' faces (condition 2). The percentage of correct answers per image pair was between 67.50 ± 13.86 % and 77.50 ± 13.63 % in condition 1 and between 64.24 ± 16.26 % and 72.80 ± 9.71 % in condition 2. For both conditions and all image pairs, a two-way repeated measures analysis of variance (ANOVA) of image pair and condition was carried out. It revealed a significant main effect of conditions, $F(1, 9) = 5.366$, $p < 0.05$, showing that the participants were answering correctly more often in condition 1, which included their own faces. The level of correct decisions was about the same for all image pairs (x-axis) as there was no effect of image pair or interaction between image pair and condition indicated (Figure 6).

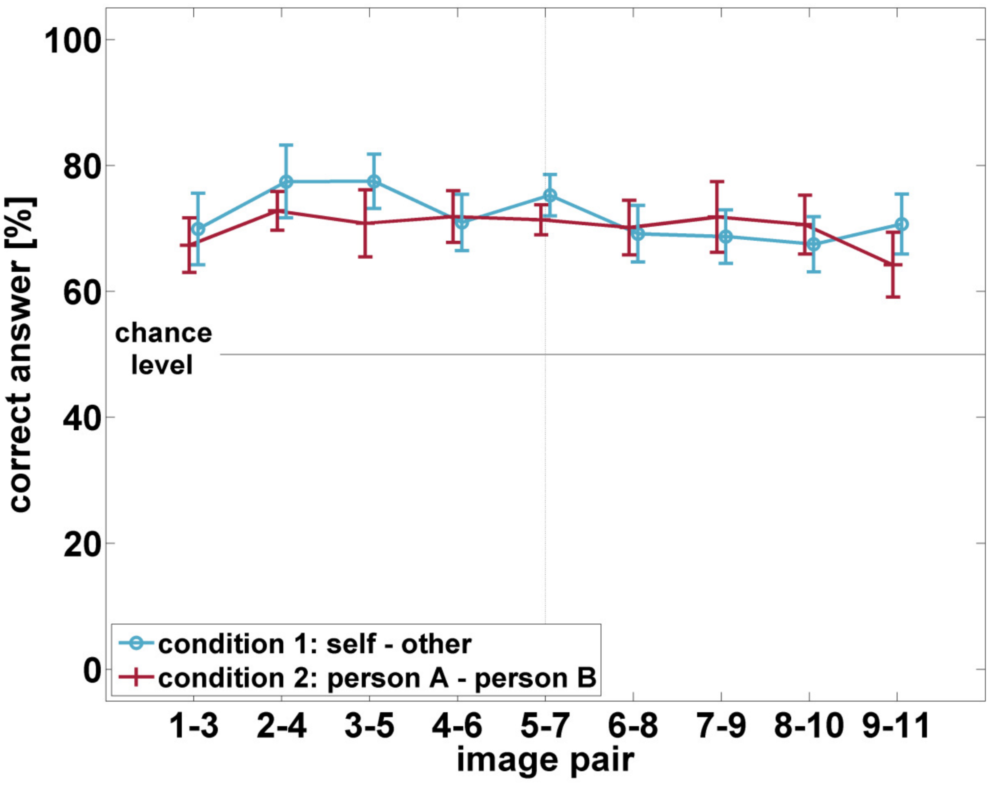

Figure 6. Results of the two discrimination task conditions. The graphs show the average percentage of correct answers with standard error of the mean (SEM) for each image pair that the participants had to discriminate between. The average percentage was about 72 % in condition 1 (blue line) and about 70 % in condition 2 (red line). The x-axis shows the image pairs, the y-axis the percentage of correct answers.

In addition to the percentage of correct answers, also the average response times were investigated. Average response times of all participants and conditions displayed similar results (Figure 7). In condition 1 the response times for the image pairs were between 2.09 ± 0.50 s and 2.40 ± 1.13 s, while in condition 2 they were between 2.14 ± 0.57 s and 2.25 ± 0.64 s. On average, participants needed 2.24 ± 0.12 s per trial in condition 1 and 2.20 ± 0.03 s in condition 2. A two-way repeated measures analysis of variance (ANOVA) indicated no differences between the conditions or image pairs and also no interaction between these factors.

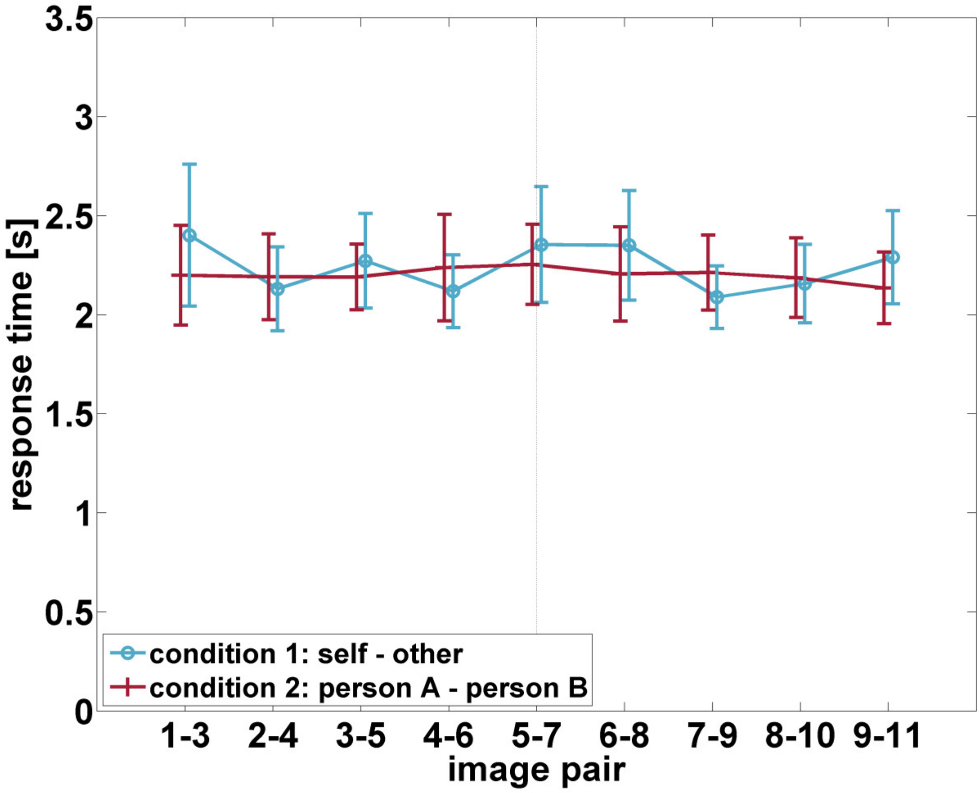

Figure 7. Average response times of the two discrimination task conditions. The graphs show the average response times with standard error of the mean (SEM) for each image pair the participants had to discriminate between. In both conditions response times were about the same across image pairs. The x-axis shows the image pairs, the y-axis the average response time.

4.2.1.2. Categorization task

Response curves were fitted with psychometric functions (error function) for each participant

$$erf(x) = \frac{1}{\sqrt{\pi}} \int_{-x}^{x} e^{-t^2}\, dt \qquad (1)$$

which were adjusted to run between zero and one

$$f(x) = \frac{1}{2} erf(ax - b) + \frac{1}{2} \qquad (2)$$

Individual PSE-levels of all participants were determined with this function (2). A paired t-test showed no difference of the PSE-levels between condition 1 and 2. However, an F-test for equal variances at the PSE-level revealed a significant difference between the conditions, $F(9, 9) = 4.552$, $p < 0.05$, with condition 1 having a greater variance. This indicates that showing (morphed) pictures of oneself led to a different answering behavior than just pictures of other people.

In Figure 8 the average decisions of all participants are shown as a function of morph-level. At high morph-levels (large amount of “self” or person A, left side of figure), participants pressed the “self” / “person A” key most of the time. Low morph-levels and therefore little amount of “self” / “person A” (right sides of figure) resulted in more keystrokes associated with “other” / “person B”. Both functions, based on the averages of all participants, crossed the 50 % decision level (y-axis) near the 50 % morph-level (x-axis), indicating a point of subjective equality (PSE) which matches the physical “mean” of the morph series.

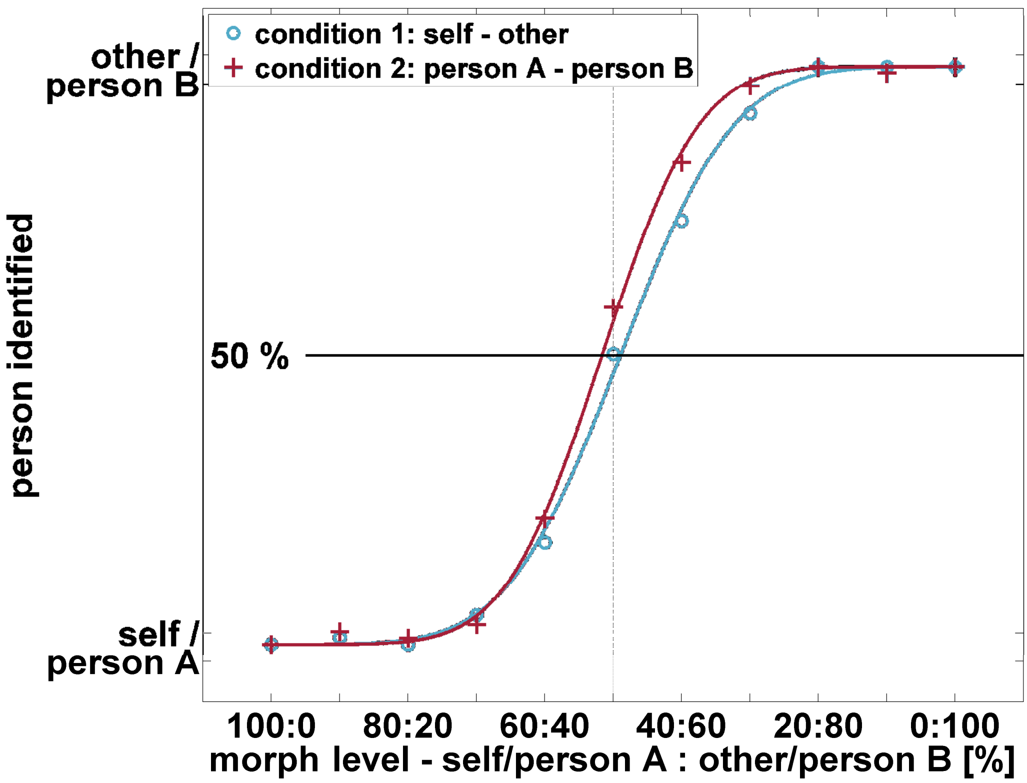

Figure 8. Results of the two categorization task conditions. The markers (+ and o) show the average answers of all participants at the given morph-level with psychometric curves fitted to these data. In both conditions the 50 % decision level was reached near the 50 % morph-level. The x-axis shows the morph-level with the ratio of self to other person (blue curve) and person A to person B (red curve). The y-axis shows which person was identified by the participants (self or other and person A or person B, respectively).

Analogous to the discrimination task, response times of the categorization task were analyzed, too. The average response times for both conditions were similar: 2.54 ± 0.46 s in condition 1 and 2.46 ± 0.26 s in condition 2. In condition 1, participants' response times were between 2.26 ± 0.24 s and 3.02 ± 1.04 s. In condition 2, their response times were between 2.21 ± 0.15 s and 3.01 ± 0.72 s. A two-way repeated measures ANOVA of morph-level and condition showed a significant main effect of morph-level, $F(10, 90) = 6.834$; $p < 0.001$, but not of condition or any interaction. A clear dichotomy in both conditions was present: At high and low morph-levels response times were lower than at morph-levels around 50 % as pairwise comparisons of pooled data revealed (Figure 9 & appendix Table 2).

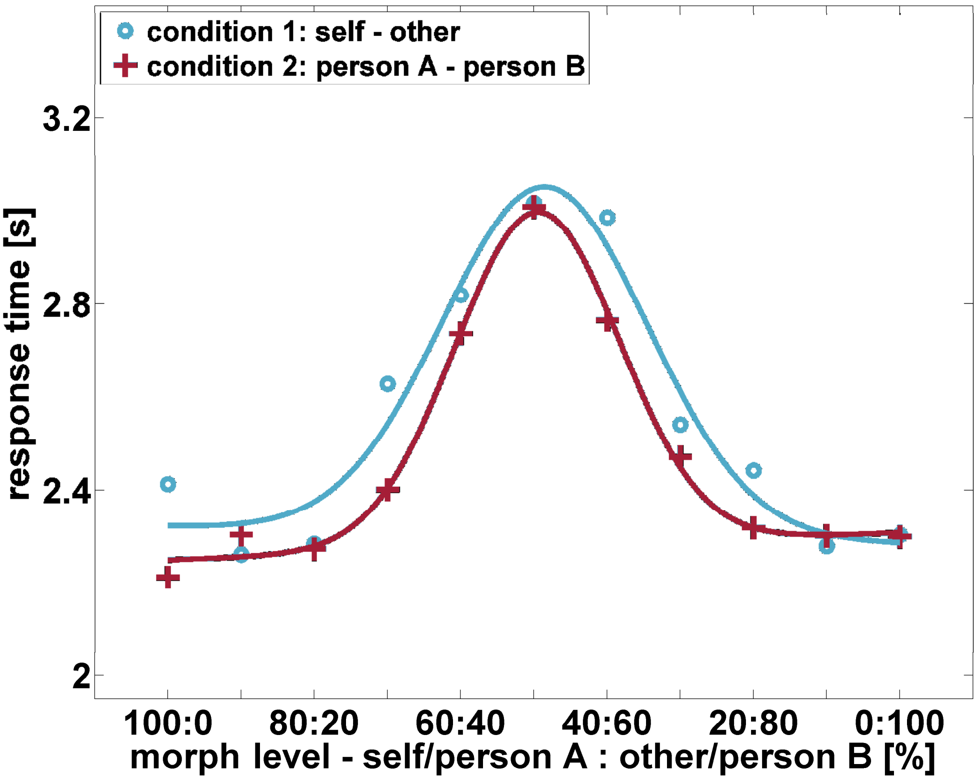

Figure 9. Average response times of the two categorization task conditions. Each marker (+ and o) shows the average response time at that morph-level of the respective condition. The graphs were fitted combinations of a linear and Gauss distribution for each condition. Response times in both conditions were comparable and were larger around 50 % morph-level. The x-axis shows the morph-level, the y-axis the average response time.

4.3. Discussion

In this experiment participants were first asked to discriminate between pictures of faces that had a dissimilarity of 20 % as well as identifying persons from sequences of morphed pictures, both with themselves being part of the pictures and other people only (conditions 1 and 2). Participants were equally good in discriminating pictures no matter what image pair (morph-level) was shown, which was contrary to the hypothesis. However, a significant difference between conditions was found. Image pairs that contained the face of the respective participant were more often correctly distinguished. Although no significant interaction was present, one could get the impression that in condition 1 for the first half of image pairs (1-3, 2-4, 3-5) the correct answer rate is greater than for the second half, which lies at about the same level as in condition 2 (Figure 6).
The amount of the participant's face in the morph picture is greater for the first image pairs. Therefore, discriminating between two images, e.g. 2 and 4 or 3 and 5, could result in a perceived discrimination like "the third picture looked more like 'me'" or "the third picture looked more like 'the other person'" as these image pairs contain more or less amount of "self", while at the same time the total amount of "self" is still large in both pictures. For image pairs like 7 and 9, the difference between the amount of "self" is still the same, but since the total amount is quite small in both pictures, the participant rather distinguishes between two pictures of "other" people because there might be not enough information in them to attribute any of it to himself, resulting in lower discrimination performance.

Here in experiment 1, no increased discrimination performance could be found around the 50 % level (that is image-pair 5-7). This is in contrast to the results of Beale & Keil (1995) who found an increased discrimination rate at the same morph-level as where in categorization the 50 % identification level was crossed, that is at the same PSE-level. However, they used morph sequences of famous people as stimuli. It is known that discrimination between two items is easier if they cross the category boundary than if they belong to the same category, like in colors (Bornstein & Korda, 1984), in phonemes (e.g. Liberman et al., 1957) and as Beale & Keil found, in faces. Beale & Keil's interpretation was that "categorical perception effects can be acquired through experience", at least in the case of faces, since morph sequences of faces do not naturally exist. However, they were unable to find the effect in pictures of nonfamous people in that study.
In contrast to the study by Beale & Keil (1995), and in line with experiment 1 of this thesis, Bülthoff & Newell (2004) did not find an increased discrimination rate in their study about sex judgements in morphed faces.

In the discrimination task of experiment 1, no significant in- or decrease in correct discriminations along the morph-level and image pairs could be found, neither in condition 1 nor in condition 2. Though, the average percentage of correct discriminations was larger in condition 1. It could be that the participants' lack of familiarity with the faces of the other persons prevented categorical perception effects in discrimination, although a repeated presentation - as it happens during the experiment - should lead to familiarization. Some participants also verbally reported at the end of the experiment that they "wouldn't get the faces out of their minds for some time". This is in accord with Levin & Beale's argument from the year 2000 that familiarization during an experiment can be sufficient for categorical perception effects to occur.
The unknown level of familiarization could indeed explain this result: Since people should know their own face well enough, discrimination performance between pictures could be increased if the amount of "self" was large enough (see above). If the amount of "self" in the morphed

pictures dropped, and therefore the amount of “other person” was raised, the performance could decrease (here only visually indicated in the figure). In condition 2 where only “other persons” were displayed, there was no difference in discrimination performance, also not visually.

Discriminating between pictures, be it people, scenes or patterns, can also be accomplished by computers. It does not necessarily involve an identification process or holistic perception of or familiarity with the displayed picture, but it can be done purely be comparing the pictures, e.g. on a pixel by pixel basis of their grey/color values. The results of the discrimination task are thus a bit ambiguous since no clear evidence for (or against) a categorical perception could be found (Figure 6).

In the categorization task, participants were asked to judge if a (morphed) picture of a person looked more like one or the other person. If there was no categorical perception, one would expect a straight diagonal line leading through the morph-levels since the response behavior would directly reflect the amount of morphed picture; e.g. at 30 % morph-level responses should be at 30 %, too. The stronger the categorical perception is and the sharper the categorical boundaries are, the steeper the psychological function should be - becoming a Heaviside step function at the extreme, where there is no smooth transition from one to the other category anymore.
Here, the response curves are considerably different from a straight line, confirming a categorical perception as it was expected for faces (Beale & Keil, 1995) and facial expressions (Etcoff & Magee, 1992; Calder et al., 1996 b). For both conditions, the PSE-level was about the same, though, the variance at this level was greater in condition 1 (self - other), meaning a less uniform boundary between these face categories. This suggests that self-perception (in morph sequences) varies between participants while the perception of other persons is more equal.
Verbal reports of some participants indicated different strategies to solve the task. Some were looking at the impression of entire face, while others were focusing on certain features, e.g. eyes, and if these looked like their own ones.

In summary, no increased discrimination accuracy could be found at the category boundary with the given familiarity level, although, categorical perception took place as psychometric response curves in the categorization task revealed. Discrimination performance was increased for faces of the participants themselves; so was the variance at the PSE-level of the categorization task. These results show that one’s own face is well represented in memory. Furthermore, in the analysis of response curves in categorization tasks, individual face pairings in morph sequences need to be considered since PSE-levels vary depending on the pairing, especially when participants’ own faces are present in the morph sequence. Since no increased discrimination performance could be found, familiarization prior to the experiments is advised.

To further investigate the perception of morph sequences and the categorical boundary between two categories and faces, experiment 2 was designed and carried out. Moreover, the representation and the characteristics of faces in these morph sequences were investigated.

5. Experiment 2 - Face-space

All the different faces one is familiar with need to be memorized and represented in the brain. One strategy how to represent all these faces would be to code the differences of a certain face from a “prototype face”, which could be thought of as an “average face” (Valentine, 1991). Faces coded that way would create a high-dimensional “face-space” in which all variants of a face, e.g. size of the nose, eye color etc. are individually represented (Valentine, 1991). A certain familiar person’s face would be a unique point in this multi-dimensional space of face representation. The more familiar one becomes with a person, the larger and more detailed this point (or rather area) in space is supposed to become, too, since one is able to recognize the person even if lighting conditions are different or if the face is partially occluded.
Another strategy would be to represent faces in a high-dimensional “face-space”, again. However, this time without a dedicated prototype face in its center, but only the distances between familiar faces are used to span the face-space.
Faces in this face-space (with or without a prototype as anchor) are distributed anisotropically and have different distances between them. This is dependent on the faces a person is familiar with and the features of these faces. In this experiment, it was investigated how this unequal distribution and different spacing influences face representation. Two factors that are hypothesized to play a role here are distinctiveness and typicality of a face.

Several studies about face recognition were focused on distinctiveness and typicality of faces, aiming towards the question how different faces and identities are recognized and discriminated. A distinctive face has certain prominent features, e.g. “intense” eyes, “sharp” nose, squared shape, etc., that makes it stick out compared to an average or typical face. A typical face, in turn, has none of these features and has a “softer” look to it. It has been shown that typical faces are more likely to be identified as a face from jumbled faces (Valentine & Bruce, 1986; Valentine, 1991; Valentine & Endo, 1992). Such typical faces are also more likely to be rated as “attractive” (Deffenbacher et al., 1998; Rhodes et al., 2000). However, recognition memory for typical faces is also inferior to faces that are more distinctive (Light et al., 1979). Distinctiveness in faces can be seen as physical deviation from the average (Bruce et al., 1994). It has been found that distinctive faces are recognized faster and more accurately from other faces (Light et al., 1979; Bartlett et al., 1984; Valentine & Bruce, 1986; Valentine & Endo, 1992).

In this experiment, participants were asked to set the physical categorical boundary between two faces, that is, the 50 % morph-level, by themselves. Thereby, it was investigated if the unequal distribution of faces in the face-space influences the decisions, which means that the perceived categorical boundary was different from the physical categorical boundary. Such a deviation would appear as a “shift”. Finally, participants were asked to rate distinctiveness of and similarity between faces. It was hypothesized that these factors influence the perceived categorical boundary and might correlate with occurring shifts.

5.1. Material and methods

5.1.1. Participants

20 adults of European descent (ten male, ten female) aged between 18 and 33, with normal or corrected to normal vision participated in this study; most of them were undergraduate students of the University of Tübingen. Participants were informed about the collected data, the general procedure and that they were free to terminate their participation at any time. Written informed consent procedures adhere to the guidelines of the Declaration of Helsinki and was obtained from all participants.

5.1.2. Setup

Participants were placed in a distance of about 50 cm in front of a PC screen (for more details see general material and methods above). The room was dimly lit and lighting was kept constant during the experiment. Stimuli were presented with Psychtoolbox 3 in MATLAB R2012a (MathWorks, Natick, MA, USA).

Images of faces, which were presented on the screen during the experiment, covered about one third to one half of the computer screen vertically, depending on the stage of the experiment (learning task or main experiment), and were clearly visible. These faces were taken from morph sequences, but in contrast to experiment 1, here, the morph intervals were only 4 % morph-level, resulting in smoother transitions.

Color photographs of faces were also printed on white cardboards (DIN A5, 14.8 x 21.0 cm) and covered about 10 x 15 cm. These were used in the learning task (see below).

The people whose faces were shown throughout the experiment were students, too, but unknown to the participants.

5.1.3. Procedure

The experiment consisted of three blocks: The first block functioned as learning task. Here, the participants were handed over ten cardboards with different faces (5 male, 5 female) and encouraged to carefully look at and memorize them as they wanted since they should find them later among other faces. There was no time limit for this learning block, but participants needed about two to three minutes for doing that.
To assess whether they learned and memorized the faces well enough, participants were asked to pass a recognition task. For this, pictures of 20 different faces (the ten faces they should memorize earlier and ten new unknown ones, balanced for gender) were randomly presented on the computer screen. Participants should indicate by pushing one of two buttons if this was a memorized or a new face (old-new-task). They were given feedback if their decision was correct or wrong. Each of the ten previously learned faces had to be correctly identified five times to pass the test. The number of trials varied depending on participants' learning and recall performance but the block took about five minutes on average.

After passing the recognition task, the second block and main part of the experiment started. Participants were shown three faces on the monitor. The ones to the left and right were unmorphed 100 % images of the two persons of the morph sequence. The face in the middle of the screen was pseudo-randomly taken from that morph sequence (16 %, 20 %, 24 % or 76 %, 80 %, 84 %); in other words they started in the "lower" quarter or "upper" quarter of that morph sequence (Figure 10). Participants could swap the face in the middle of the screen to another one from the morph sequence by pressing the left or right arrow key on the keyboard. With every keystroke the face changed by 4 % morph-level converging towards one of the faces on the sides, depending on which key was pushed. Mathematically speaking, an oriented permutation took place. Participants were asked to push the buttons and "set the face they thought would equally consist of both persons", like turning a rotary control. When they thought they had set the 50 % morph-level, participants should push the space bar and the next trial began, starting with a fixation cross for 300 ms. There were 240 trials in total, consisting of 20 face pairs (ten male, ten female) times six starting images times two repetitions. All images were presented in random order, though, care was taken that not two images of the same morph sequence were presented in a row.

Afterwards, the third block began in which participants were asked to rate the previously used pictures once according to their similarity to each other (pairwise, in all possible combinations) on a scale of one to seven, where one meant very similar to each other and seven meant very different from each other. Subsequently, participants should rate once how distinctive each face on its own was, again on a scale of one to seven, where one meant very average and seven meant very distinctive.

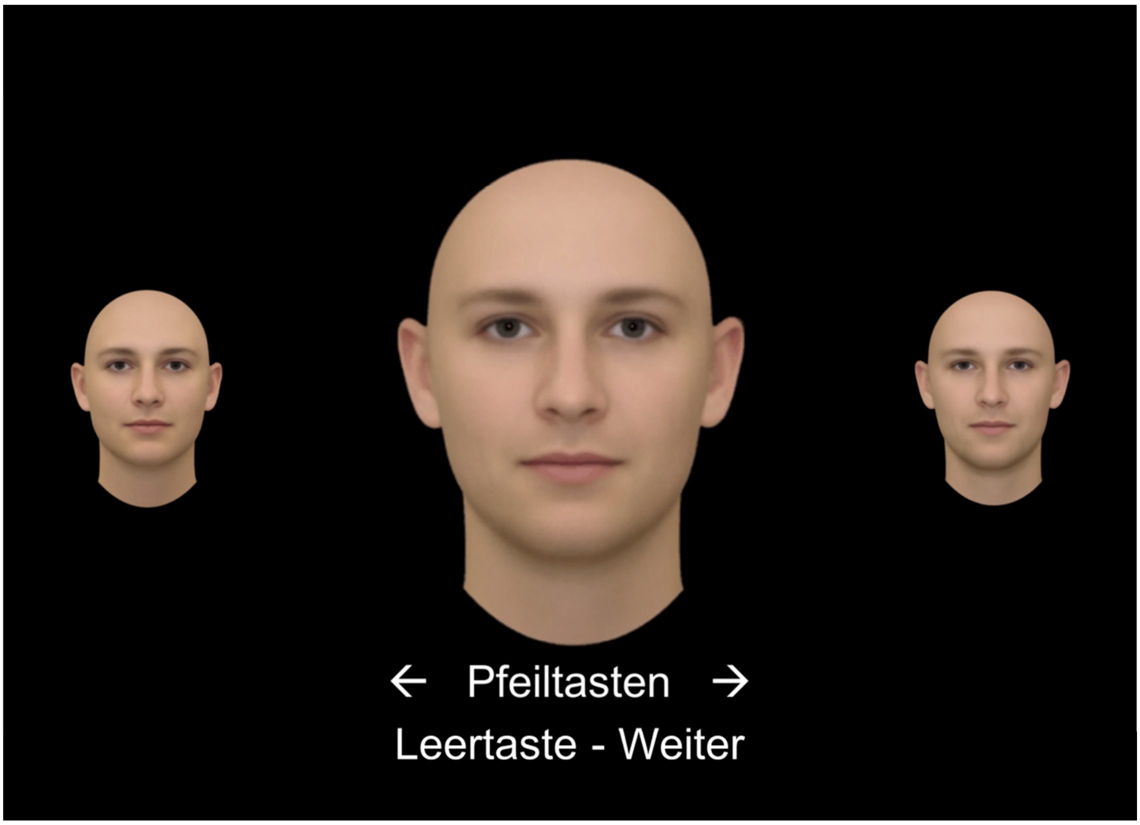

Figure 10. Example of a trial of the experiment's main part as it was displayed to the participants. To the sides of the screen, the two unmorphed faces or endpoints of the morph sequence were shown. The face in the middle of the screen was taken from the morph sequence and participants were able to swap the face to another one from that morph sequence with the left and right arrow keys on the keyboard.

5.2. Results

All participants completed the experiment as instructed. On total average, the participants set the subjective 50 % level for nine out of ten female face pairs (A to E, red, Figure 11) at the 50 % morph-level. A significant shift towards face E was found for face pair DE, $t_{DE}(19) = 3.301$; $p < 0.05$. For the ten male face pairs (V to Z, blue, Figure 11), participants set the subjective 50 % level for six out of ten face pairs at the 50 % morph-level. For five male face pairs (DE, VW, VX, WX, XZ), t-tests indicated significant shifts away from 50 % morph-level. For both face pairs VW, $t_{VW}(19) = 4.334$; $p < 0.001$, and VX, $t_{VX}(19) = 2.697$; $p < 0.05$, shifts were towards face V. For the face pairs WX, $t_{WX}(19) = 4.071$; $p < 0.001$, and XZ, $t_{XZ}(19) = 2.780$; $p < 0.05$, shifts went away from the 50 % level towards face X in both cases. For all of the 20 face pairs the subjective 50 % levels were set on average at or close to the 50 % morph-level with the largest deviation being 4.7 % for face pair VW.

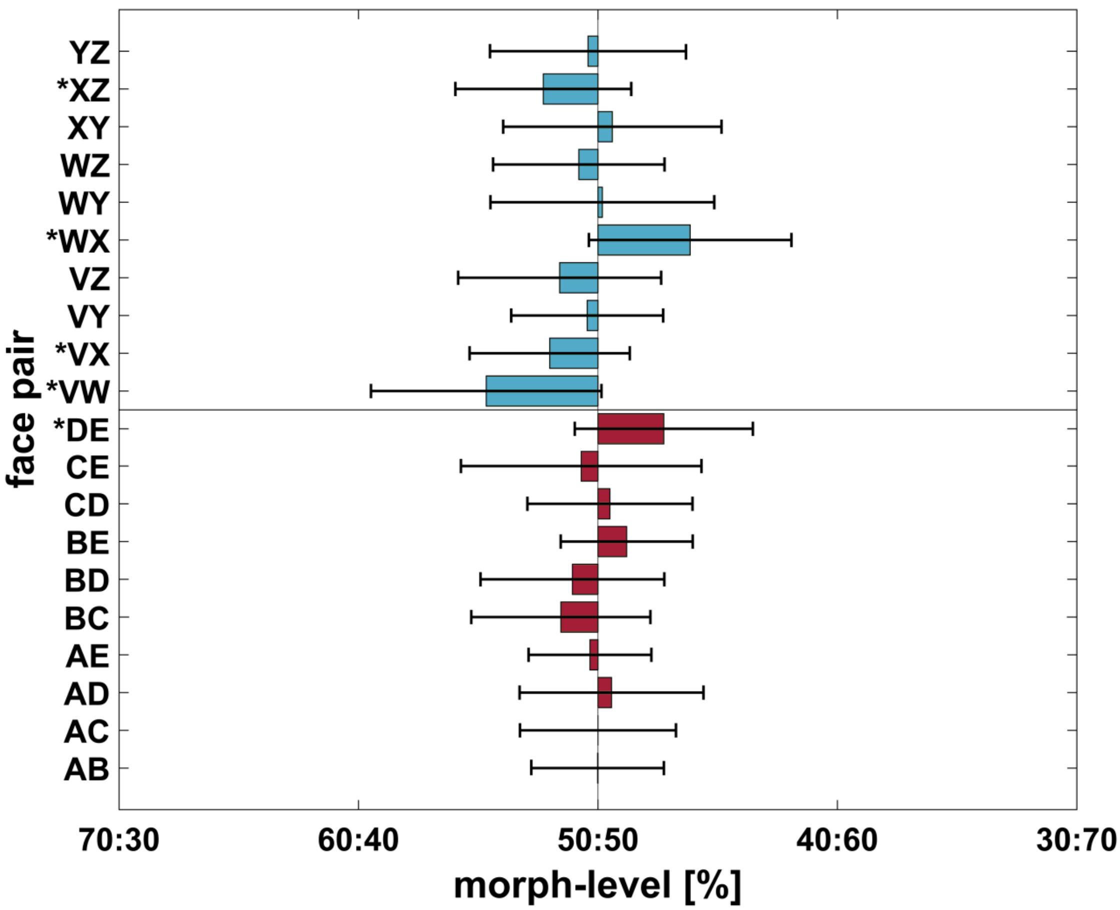

*Figure 11. Average set levels and shifts for the 20 face pairs. 15 out of 20 face pairs were set at the 50 % morph-level on average. Five pairs were significantly shifted away (tagged with a star *). The x-axis shows (a segment of) the morph-level, the y-axis the face pairs with female (A - E, red) and male (V - Z, blue) faces. Error bars show standard deviation.*

Besides the setting of the subjective morph-level, the response times and number of clicks were analyzed. The response time was the time that passed from the fixation cross disappearing to the push of the space bar key by the participant, which indicated that the subjective 50 % level was set. Number of clicks is the number of keystrokes participants needed for changing the displayed face and "moving" along the morph-level in 4 % steps until the confirmation key (space bar) was pressed. Figure 12 shows the average number of clicks and response time of all participants across trials. One can see that at the beginning of the experiment participants needed more clicks and more time to complete a trial. Later on, they got faster and needed less clicks to do so. Peaks and valleys (e.g. near trial 60) are due to individual response characteristics of the participants. For participants' individual number of clicks and response times see appendix Figure 50.

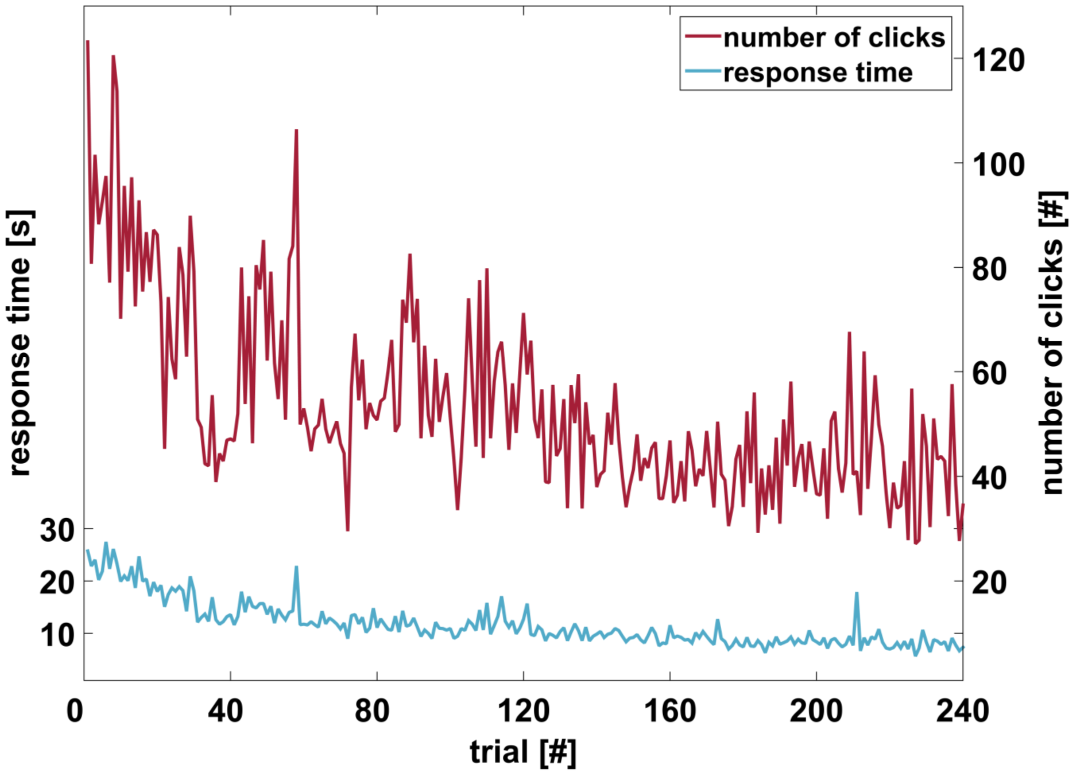

Figure 12. Mean response time and number of clicks across trials of all participants. Both response times and number of clicks decrease with time, which indicates a learning effect. The x-axis shows the trial, the y-axes show the response time (blue, left side) and number of clicks (red, right side).

The average response time and number of clicks per face pair are displayed in Figure 13 a) and b). Data was averaged over all participants and trials. There was no significant difference between female and male faces neither in response time (both 11.72 s) nor in number of clicks (female: 53.9, male: 53.0) on average. Though, a substantial variance was present for individual face pairs. Face pairs for which on average a significant shift away from the 50 % morph-level was evident in the last analysis (Figure 11) were tagged with a star in Figure 13, too. As one can see, both response time and number of clicks for those five face pairs are still comparable to the other 15 face pairs for which no significant shift away from the 50 % morph-level was found.

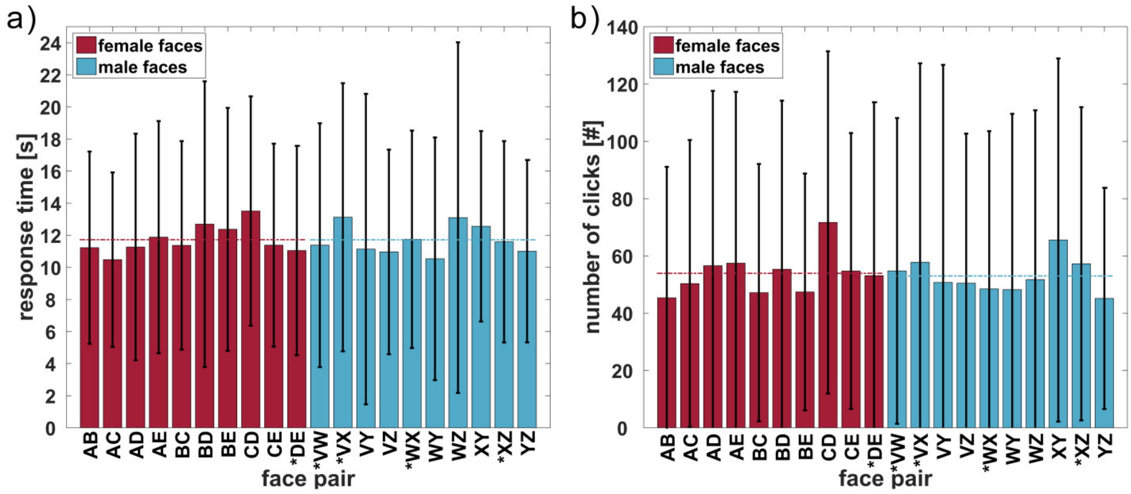

Figure 13. Response times and number of clicks. a) Average response time per face pair. For both female (red) and male (blue) face pairs the overall average response times (horizontal dotted lines) were identical with 11.72 s. b) Average number of clicks per face pair. The numbers of clicks for female (53.9) and male (53.0) face pairs were similar. Face pairs that had a significant shift away from 50 % morph-level on average in the last analysis are tagged with a star. The x-axes show the face pairs for which the subjective 50 % levels should be set, the y-axes show a) the response time and b) the number of clicks. Error bars show standard deviation.

5.2.1. Shifts in face-space

In addition to the analysis whether there was a shift away from 50 % morph-level for a certain face pair, also participants' individual direction of shifts in these two face-spaces was evaluated. This was done for male and female face pairs, each consisting of five faces (A to E and V to Z). Therefore, the mean shift direction (if there was a significant shift) per face pair and participant was counted and summed up. For example, participant No. 3 set the subjective 50 % level of face pair AC on (his personal) average at 56 % and that individual shift was significant, tested with a t-test (12 repetitions per face pair, see above). This would count as one time shift towards face C away from face A (see appendix Table 3 for t-tests). This procedure was done for all participants and face pairs, resulting in a total number of 28 significant shifts for female faces and 53 significant shifts for male faces. In Figure 14, the relationships and shifts between the faces are illustrated. The numbers that count shifts away from a face and towards a face are twofold. This means, for example, that a single shift from face C to A was counted both as "away from C" and "towards A". For each face (A to E and V to Z), the total numbers of cases with shifts towards (positive sign) and away (negative sign) from that face were summed up, resulting in a single arrow pointing into the direction in which more shifts occurred. The bigger the arrow towards a face, the more often shifts occurred in that direction.

There were faces like the male face "V" that acted like an attractor: There was a large amount of shifts towards that face. On the other hand, there were faces like "W", "Z" and "D" that acted like repellers - shifts went away from these faces. In general, these effects were stronger for male than female faces. For male faces, a total number of 53 shifts occurred, resulting in 106 counted directions in the face-space diagram. For female faces there were 28 shifts, resulting in 56 counted directions.

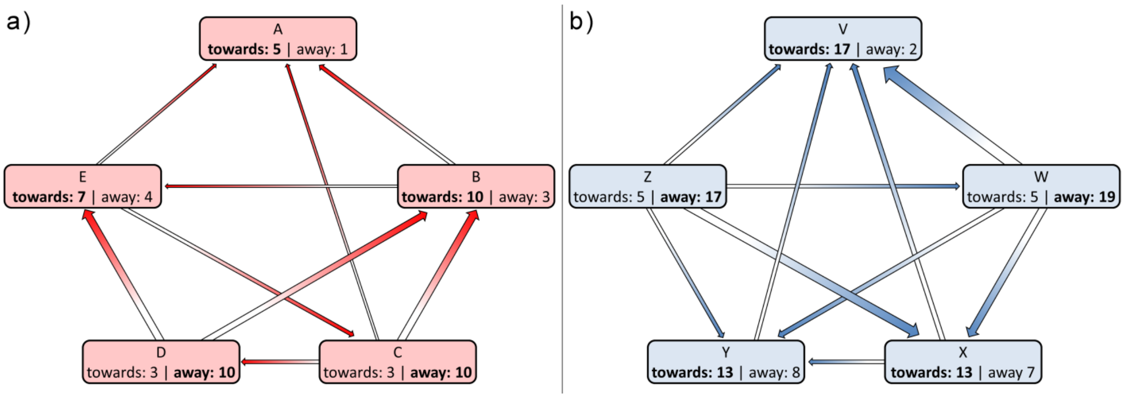

Figure 14. Shifts in face-space. Significant shifts towards or away from a) female faces (red) and b) male faces (blue) are displayed. There were more shifts for male faces than female faces. Numbers indicate the total amount of shifts going towards or away from the respective face. Predominant directions are written in bold. Arrows show net shifts into a certain direction; the bigger the arrow, the more shifts in this direction occurred.

Two typical factors "distinctiveness" and "similarity between two faces" were further investigated as they could provide explanations why faces "attract or repel" shifts. Participants were asked to rate every face in terms of distinctiveness from 1 (very average) to 7 (very distinctive). The results of this rating are displayed in Figure 15.

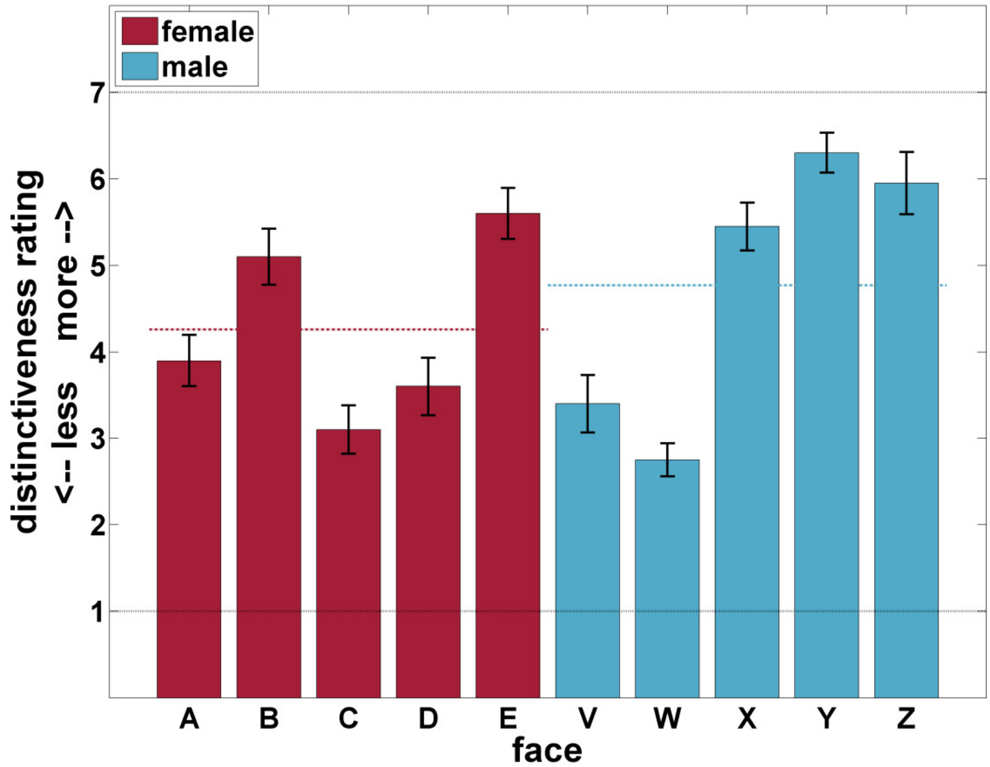

Figure 15. Distinctiveness ratings on the presented faces. Male faces (blue) had a higher rating on average (dotted lines) than female faces (red). Also the difference between high and low ratings within a group was larger for male faces. The x-axis shows the rated faces, the y-axis shows the rating between 1 (very average) and 7 (very distinctive). Error bars show standard deviation.

On average, the distinctiveness rating for female faces was 4.26 ± 0.71 and 4.77 ± 0.57 for male ones with a significant difference between these two groups (Mann-Whitney U = 9239.5; $n_1 = n_2 = 100$; $p < 0.05$). Female faces were rated closer to the mean of the scale (4), whereas male faces were rated more distinctive on average. Still, in both groups there were faces that were rated rather indistinctive (C and D as well as V and W) and others that were quite distinctive (B and E as well as Y and Z).

Further, the similarity between face pairs had to be rated by the participants. Results are shown in Figure 16.

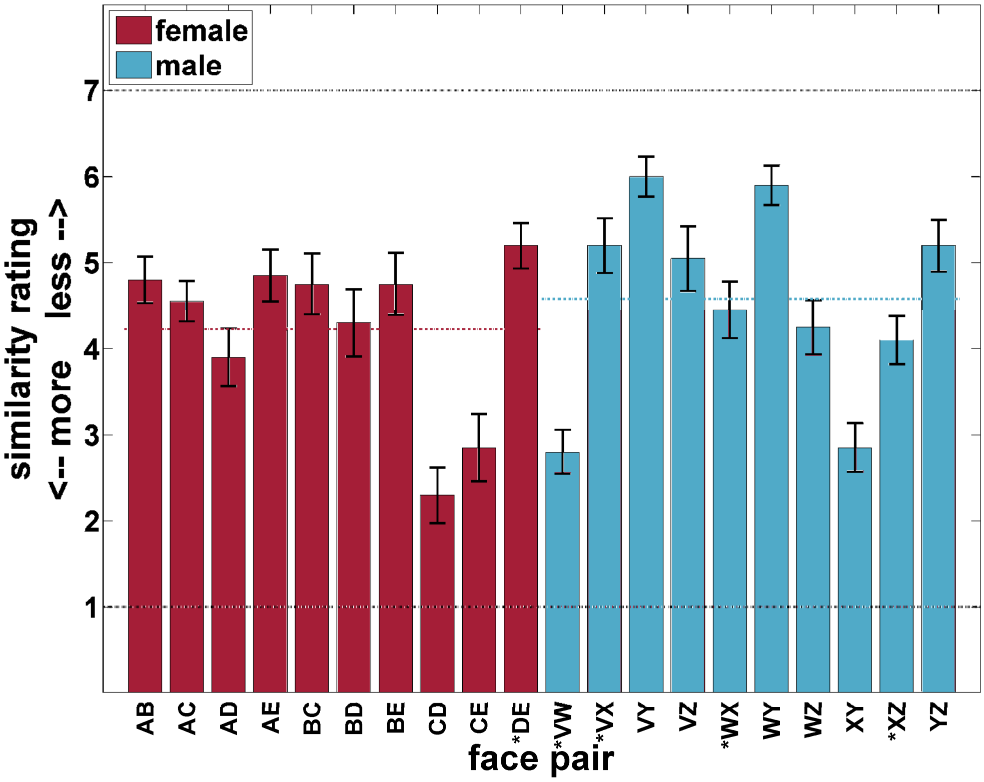

Figure 16. Similarity ratings between faces. Female face pairs (red) were rated more similar on average (dotted lines) than male face pairs (blue). Ratings of female face pairs deviated less from the average rating than male face pairs. Stars indicate face pairs that had significant shifts away from 50 % morph-level in the first analysis (see above). The x-axis shows the rated face pairs, the y-axis shows the rating between 1 (very similar) and 7 (very different). Error bars show standard deviation.

Seven of the ten female face pairs were rated to be averagely similar (close to a rating of 4). Also six of the ten male face pairs had an average similarity. Though, some of the male face pairs had a large dissimilarity (VY and WY). On the other hand, some face pairs were rated to be very similar (CD and CE as well as VW and XY). Face pairs that had significant shifts in morph-level were both quite similar (VW) and average (WX) as well as rather dissimilar (DE and VX), which means no indicative pattern was present here. On average, female faces were rated as 4.23 ± 0.67 and male faces as 4.58 ± 0.71. There was no significant difference in similarity rating between these two groups (Mann-Whitney U = 38021; $n_1 = n_2 = 200$; $p > 0.05$).

To investigate whether there was a correlation between shifts and distinctiveness as well as between shifts and similarity, data were compared with respect to these properties and results are displayed in Figure 17.

In the experiment, participants were rating distinctiveness individually per face and did not compare it to a second face as it was done for the similarity rating between two faces. Because of that, distinctiveness ratings between all faces were calculated and subsequently compared with the respective shifts between these faces. For instance, for the face pair CE there was a shift towards C, which means a negative shift number since it was towards the first face (C) in this notation (see also Figure 11). Further, the face C was rated less distinctive than face E (Figure 15), which leads to a negative number in distinctive difference (C - E). This results in a data point in the third (lower left) quadrant of Figure 17 a).
For female faces, no correlation between distinctiveness and shift was present, $r_S(198) = -0.104$; $p = 0.142$. For male faces, a trend was observed, $r_S(198) = -0.130$; $p = 0.066$, indicating that shifts occurred towards the more distinctive face. For pooled data (male and female faces) there was no correlation.

For the correlation between similarity of face pairs and shifts, absolute numbers were used for the shifts. This is because similarity ratings between faces were absent of sign (e.g. similarity rating of face A to B was 4.8, while rating of B to A was 4.8, too).
The similarity of face pairs and shifts displayed a different result then: For male faces there was no correlation at all, $r_S(198)= -0.014$; $p = 0.841$. However, for female faces a negative correlation has been found, $r_S(198)= -0.142$; $p < 0.05$. The more similar faces were, the larger the shift was. Again, no correlation was present for pooled data.

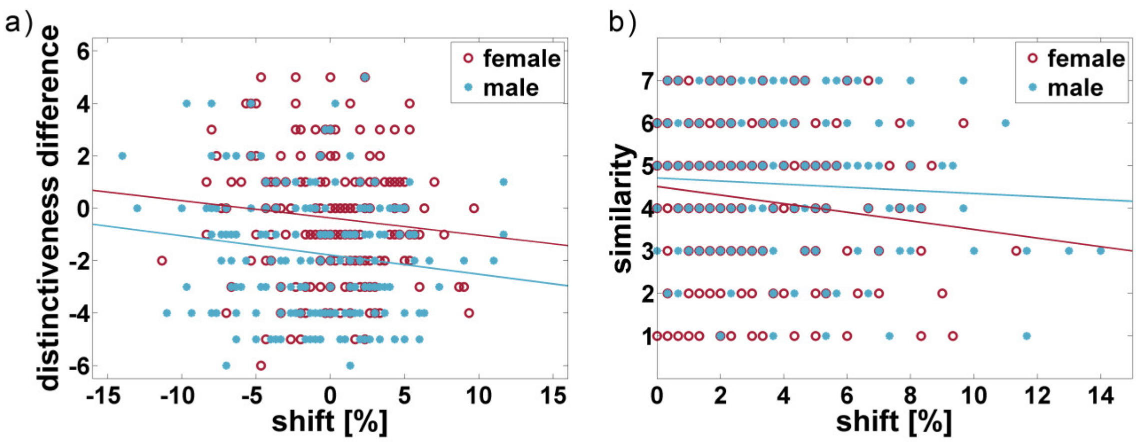

Figure 17. Correlations between shifts and distinctiveness and similarity. a) Distinctiveness differences were uncorrelated to the shifts in morph-level for female faces. A trend was observed for male faces where shifts occurred towards the more distinctive face (negative y-axis value). b) Similarity ratings were uncorrelated from shifts in morph-level for male faces; however, a correlation was present for female faces. Each marker indicates one rating per face pair and participant. The x-axes show shifts in morph-level, the y-axes show the difference in distinctiveness ratings and similarity ratings.

These correlation analyses were also done for just those five face pairs that showed significant shifts in morph-level on average (see Figure 11). This led to an enhancement of the previous results for shifts and distinctiveness: No correlation for female faces was found, though, a strong correlation between distinctiveness difference and shifts for male faces, $r_S(78) = -0.297$; $p < 0.01$, towards the more distinct face was observed (see appendix Figure 51).
For the comparison between shifts and similarity of face pairs, no correlations could be found for female or male faces.

5.3. Discussion

In experiment 2, participants were asked to set the physical categorical boundary (= 50 % morph-level) of morph sequences and rate faces and face pairs by their distinctiveness and similarity. Participants set this individual 50 % morph-level to a large amount quite well and only for few face pairs a small but significant shift away from this 50 %-level was present. It is striking that only for one female face pair such a shift emerged, while for male face pairs this was present in four out of ten times. On an individual subject-level, shifts occurred more often, but they also varied more often between participants. Figure 14 provides a detailed view on the shifts between the face pairs. Some faces "attract" shifts while others "repel" them.

To investigate these differences between female and male faces in terms of shift frequency and to find reasons for these shifts in general, a closer look at the other parameters is necessary. Both response times and number of clicks were evaluated. It was found that response times as well as number of clicks decreased across trials. This means participants were faster in setting the individual 50 % morph level and needed less clicks to do so. Though, neither response times nor number of clicks provided any evidence for an influence on the occurring shift. Both factors were also very similar for all face pairs.
The other two factors that were recorded and theorized to have an influence on the shifts were distinctiveness of a face and the similarity between a pair of faces. Distinctiveness ratings provided a first evidence for the difference between female and male faces. Male faces - on average - were rated to be more distinctive than female faces. Also the similarity between male face pairs was less on average compared to female ones. This is in accordance with other findings about face characteristics where male faces typically look differently from female faces with more distinctive features (Enlow, 1990). Also Little & Hancock (2002) found that morphing or averaging multiple faces into one resulted in a more average or typical face, which was rated to be more attractive but also less masculine and less distinctive. Still, two of the male faces used in this experiment were rated less distinctive than some female faces, which can be explained by natural variance.
The average distinctiveness both for female and male faces was about 4.5 with values between 2.5 and 6.5. These numbers are in line with Burton & Vokey (1998) who explained that typicality ratings of faces are normally distributed with just few faces rated as very typical or very distinctive. However, in this experiment here, participants were asked to rate "distinctiveness". While "distinctive" does not necessarily mean the direct opposite of "typical", they are surely connected.
It is also noticeable that those faces that were rated less distinctive also got larger similarity ratings among each other (CD and VW). For faces that were more distinctive, such a correlation with others was not apparent except for face pair XY. Although pair CE received rather different ratings in terms of distinctiveness, it was rated to be quite similar to each other. It is possible that

the distinctiveness of a face can be caused by just a few features (e.g. a big nose) while other aspects of the face are quite "average". Thus it is possible that this face is perceived to be similar to another one which is rated less distinctive despite that seeming discrepancy.

It is known that faces that are less distinctive or closer to the/a facial prototype(s) (see below) are more likely to be falsely recognized in terms of identity (Light et al, 1979) but also more easily recognized 'as a face' in the first place (e.g. Valentine & Bruce, 1986). However, it is unlikely that confusion of identity took place in the experiment because it was only a small set of faces the participants were asked to rate, and they were also familiarized with them prior to the experiment. It is debated if there is a single face prototype in the center of the face-space (Valentine & Bruce, 1986; Valentine, 1991) or if there are multiple ones especially for bimodal features like gender (Baudouin & Gallay, 2006; Baudouin & Brochard, 2011). Anyway, the more distinctive features a face has, the further away it is from the (respective) prototype(s).

The findings of this experiment fit well into the notion of a multi-dimensional "face-space" where all faces are located in. C, D, V and W are less distinctive and therefore lack features which would place them further apart in the face-space. X and Y, however, are both distinctive, yet their striking features seem to be similar, resulting in a respective similarity rating. Faces C and E probably share certain features, while E also possesses some that are more distinctive. In a multidimensional space C and E could be close in certain planes while being apart in other ones.

Looking at the relation between distinctiveness and morph-level shifts, no correlation could be found if all faces were considered, though, a trend was present for male faces, showing that shifts occurred towards the more distinctive face. This trend turned into a significant correlation if only those faces were considered that had a significant shift away from the 50 % morph-level earlier.
In turn, similarity between face pairs is a predictor for morph-level shifts but only for female faces. Shifts occurred more often if similarity between (female) faces was high. Since the similarity rating does not indicate the shift direction but only points to the fact that there was a shift - in combination with shifts only happening with rather similar faces - these findings indicate that it was more difficult for the participants to actually find the 50 % level here. It also explains why for male faces, being less similar, no such correlation was present.
In other studies, the resolution of the face-space or density of faces within the face-space is discussed, too. Valentine (1991) stated that it is likely that recognizing typical faces is more difficult if encoding errors are presumed since these "typical faces are more densely clustered in face-space". This means that with encoding errors taking place it is more likely "to lead to confusion of facial identity for typical faces". In this experiment, neighboring faces in the morph sequence were extremely similar to each other (differences of 4 %). Now, if the two morphed faces were already quite similar in the beginning, it was even harder for the participants to spot differences while moving along the morph sequence, which could have led to (more) shifts for female face pairs with large similarity.
The lack of correlations for female faces if only significant shifts were considered can be explained by the small number of data, since only one face pair had a significant shift (in contrast to four significant shifts for male faces).

It is further assumed that the mental face-space, in which faces are represented in the brain, can be scaled and optimized so that faces can be well represented and discriminated. This is observed in the "own-ethnicity bias" where people can discriminate members of their own ethnicity better than people of a different ethnicity (Meissner & Brigham, 2001; Gross, 2009), because they encounter those faces on a daily basis. The process for that requires perceptual learning so that the face-space can be tuned accordingly (Valentine, 1991; Valentine & Endo, 1992; Valentine et al., 2016). However, it requires more than just simple visual exposure to these faces to actually learn and encode these faces in order to tune the face-space (Chiroro & Valentine, 1995).

The results of this experiment indicate that, at least for faces that were rated similar, the discrimination or resolution limits of the participants' face-space were approached as they were not able to set the 50 % morph-level as precise as they did for less similar faces. Yet, according to the theories described above, participants should be able to increase their performance if they were trained with these faces and interacted with the people regularly.

Altogether, characteristics of faces and face representation found in this experiment are in line with results of other studies. Small shifts were found for certain face pairs indicating that different distances between faces in the face-space lead to altered perceived categorical boundaries. Shifts increased as current resolution limits of the participants' face-space were approached; or perhaps individual, but yet unknown, differences in face pairs were effecting such shifts, since distinctiveness and similarity were able to explain (only) some of those.

Neither response times nor number of clicks indicated "problematic" faces or face pairs in terms of standing out from the other stimuli. The stimuli used here seem to reflect an adequate selection from the face-space. In experiment 3, these faces were used as stimuli for priming and recognition experiments. Also shifts were investigated further; namely, whether they can be induced artificially via priming.

6. Experiment 3 - Primed face recognition

In the previous chapter, it has been shown that the anisotropy of the represented faces in the face-space leads to shifts in perceived category boundaries. In this chapter, it was investigated whether shifts can be induced artificially and thereby altering the participants' perceived identity of people's faces, including their own one.

Therefore, two experiments were designed and conducted with direct identity priming of self and other faces to gain further knowledge especially in two aspects: First, the recognition of one's own face, e.g. on a picture. To do this, the visual input has to be compared to an internally stored representation or image of oneself. In experiment 3 a), the effects of visual priming with one's own face or the face of another person on response times and response behavior were investigated. Second, the connection between one's name and image. In experiment 3 b), the effect of names as primes was in focus. Different studies reported different results on this topic (see above) where some found a priming effect of names and some did not. Therefore, in experiment 3 b) the effect of a person's name in combination with visual priming like in experiment 3 a) was explored. In both experiments masking effects (Dehaene et al., 2001) of short stimulus onset asynchronies as in Pannese & Hirsch (2010), brief prime duration and target presentation were used. Such a masking effect occurs when a stimulus is presented very briefly and accompanied by a second stimulus that is close in temporal and spatial regards. Participants were asked to identify the person shown on the test picture (= target) by pressing one of two associated keys in a yes-no-paradigm. It is hypothesized that participants respond faster if prime image and test image are similar. Furthermore, participants are expected to identify morph images as a certain person even if this person contributes to that image less than 50 % as long as prime images amplify that identity. This would mean shifts of perceived identity can be artificially induced. An account on the results of the experiments will be given in the discussion as well as a priming model presented that concludes these findings.
The terms "prime" and "target" are commonly used in priming studies to describe the two main stimuli that influence processing (prime) and which participants should respond to (target). Since in experiment 3 b) the term "target person" was used during the experiment, and to avoid misunderstandings, the target stimulus (here pictures) will be referred to as "test" picture in both experiments 3 a) and b).

6.1. Experiment 3 a)

6.1.1. Material and methods

6.1.1.1. Participants

14 adults of European descent (six male, eight female) aged between 20 and 36 (mean age 26.1), with normal or corrected to normal vision participated in this study. They were paid a fixed amount of 8 € per hour. Similar to the experiments before, care has been taken to select men without facial hair and women with no apparent facial make-up since morphing would have led to unrealistic pictures. Two of the participants were excluded from analysis as they declined to

participate in the second session of the experiment (see procedure). Participants were informed about the collected data, the general procedure and that they were free to terminate their participation at any time. Written informed consent procedures adhere to the guidelines of the Declaration of Helsinki and were obtained from all participants.

6.1.1.2. Stimuli

Twenty high quality portrait photographs of the fourteen participants and another six people were used in this experiment. These photographs were used as prime pictures and test pictures and covered about 6° of visual angle during presentation. If the picture of a certain person was displayed as prime and the test picture showed a different person (or morph picture with a considerable amount of that person) a flickering was noticeable upon switching between these two pictures. To ensure that participants were able to see prime and test pictures as two separate images and that they were unable to scrutinize flickering information (or the lack of it), all prime pictures were made slightly larger and flipped left to right so that same prime and test pictures would still result in the same visual perception as different pairings did, which means a flickering of the screen due to picture change.

6.1.1.3. Setup

Participants were placed in front of a PC screen (see general material and methods above) in a distance of about 50 cm in a dimly lit room. The room lighting was controlled and kept constant during the experiment. Stimuli were presented with Psychtoolbox 3 in MATLAB R2012a (MathWorks, Natick, MA, USA).

6.1.1.4. Procedure

Full instructions were given prior to the experiment. At the beginning, participants were presented with unmorphed or “pure” 100 % images of themselves and three other persons A, B, C (same gender) their image was morphed with. Participants were encouraged to look at these pictures carefully and as long as they liked to familiarize themselves as they would be presented with morphs of these pictures later. Upon button press the main part of the experiment started.

In each trial of the main experiment, a fixation cross was presented in the middle of the screen for 2-3 s (randomly chosen to avoid anticipation and a rhythmic response behavior of the participants). Subsequently, a prime image was shown for 30 ms, which was a periliminal stimulus at that given duration (Pannese & Hirsch, 2010). The prime image was always a 100 % picture of either the participant or one of the other three persons. Then, the test picture was presented for 170 ms, which was an image taken from the appropriate morph sequence. Afterwards, the screen went black. Participants were asked to identify the person in the shown test picture as self or other and press one of the two according buttons on the keyboard in a yes-no-paradigm (Figure 18).

Each prime-test combination (2 primes x 3 combinations) was tested at least five times at each morph-level (11 morph-levels, 10 % interval) with five repetitions at 0:100 and 100:0 percent morph-levels up to seven repetitions at 50:50 % morph-level. Stimuli were randomly selected from this set. Additionally, a real-time evaluation of the responses was carried out to determine additional morph-levels for testing. The procedure to calculate the test levels was based on parameter estimation for sequential testing (“best PEST”, Pentland, 1980; Lieberman & Pentland, 1982). Both methods were used here to gain more data around the inflection point of the psychological function which is supposed to describe the participants' answering behavior for such a task with categorical perception. There were 580 trials in total for each participant to complete. These were split into two sessions with approximately one hour each and carried out on separate days.

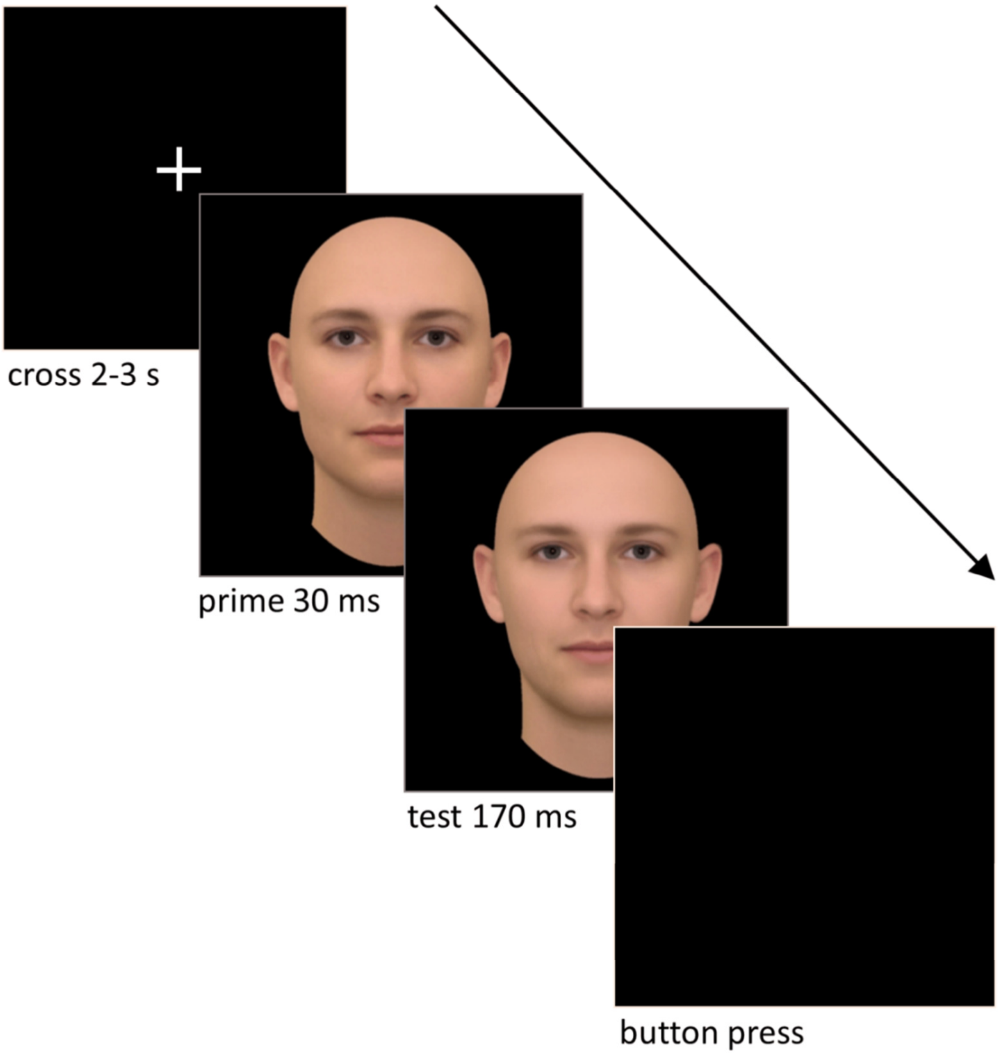

Figure 18. Sequence of a trial. Each trial started with a fixation cross for 2-3 s randomly chosen. Afterwards, the prime picture appeared for 30 ms. Then the test picture was displayed for 170 ms. After that the screen went black. The participant should indicate with a button press whether the shown test picture was him-/herself or not.

6.1.2. Results

Out of the 14 participants, two did not participate in the second session of the experiment and were therefore excluded from analysis. Also, trials with response times above 2 s were discarded (35 out of 6960). No lower limit was introduced as the fastest response was 297 ms, which is very fast but still plausible.

6.1.2.1. Response times

The average response time across all trials and participants was 770.2 ± 107.2 ms. Fast responses of about 730 ms occurred at low and high morph-levels. In these cases the test images were either "pure" pictures of the participant him-/herself or the other person. At these morph-levels the ambiguity of the person's identity was low. In Figure 19 a) - c) typical response time curves of three participants (2, 4, and 11) are shown (see appendix Figure 52 to Figure 55 for all participants). It can be seen that the response time courses differed between participants. However, all courses featured an increased response time around the center of the graph. This peak, in turn, varied in its shape (broad or pointed) and its location along the x-axis, depending on the participants, the morph-sequence (yellow, red, blue) and the prime (self, other). In Figure 19 d), the average response times of all twelve participants are shown (red and blue markers), and data were fitted with combinations of a linear and Gauss distribution. There was a peak in response times around 50 % morph-level with a maximum of 865 ms on average. Here, the test images became more ambiguous and participants needed more time to categorize them.

A two-way repeated measures ANOVA of prime and morph-level revealed a strong interaction between these factors, $F(2.68, 29.57) = 5.72$; $p < 0.01$, as well as significant main effects of prime, $F(1, 11) = 6.14$; $p < 0.05$, and morph-level, $F(1.63, 17.89) = 7.70$; $p < 0.01$. Since Mauchly's test indicated that the assumption of sphericity had been violated for morph-level and interaction, Greenhouse-Geisser corrected values were used here.

While response times were similar for ambiguous test images (morph-level about 50 %), a dichotomy could be seen at large and small morph-levels depending on the congruency of prime and the predominant component in the test picture. In congruent priming, i.e. if prime and test show (mostly) the same person, (blue curve left side, red curve right side; Figure 19) response times were in the order of 700 ms. In contrast, if priming was incongruent, i.e. if the prime showed the person not predominant in the test picture, t-tests indicated significant increases of response times of about 50 ms for the last three morph-levels in each case (blue curve right side, red curve left side). Priming with self or other faces showed similar response times since both curves (red and blue) were virtually symmetrical.

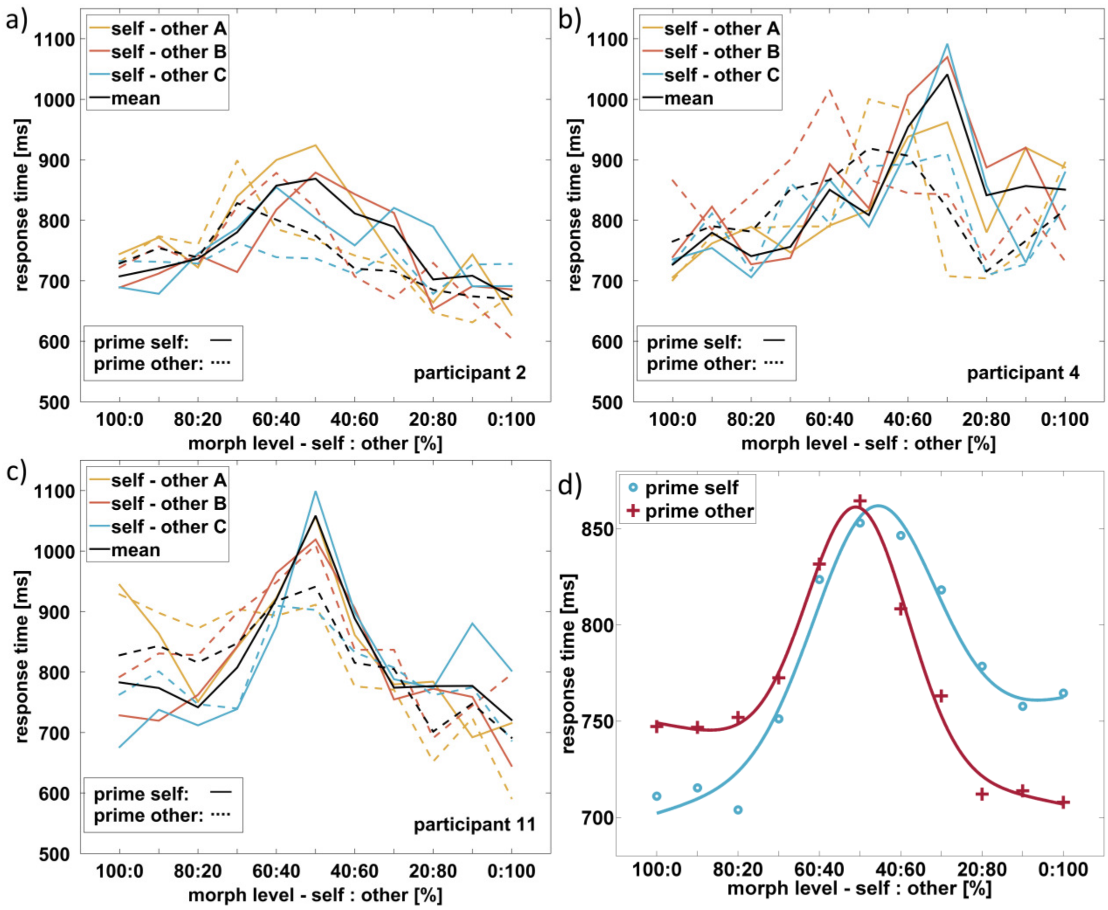

Figure 19. Response time courses. a) - c) Typical response times of three participants (2, 4, and 11) are shown. Responses were slower near the 50 % morph-level as the test image became more ambiguous. On the subject level it was difficult to distinguish between the two priming conditions "prime self" (solid lines) and "prime other" (dotted lines) for all face pairs self - A/B/C (yellow, red, blue). d) Averaged over all twelve participants and morph-sequences, a clear dichotomy was present. Responses were faster if priming was congruent with the test picture (blue curve left side, red curve right side) and slower with incongruent priming (red curve left side, blue curve right side). Each marker (+ and o) shows the average response time at that morph-level with the respective prime. The two graphs were fitted combinations of a linear and Gauss distribution for each prime. X-axes display morph-levels in percent and y-axes response times in milliseconds.

6.1.2.2. Face identification

In addition to the response time effects reported above, priming might also affect face identification. Figure 20 a) - c) shows typical psychometric functions (identification rate vs. morph-level) of three participants (2, 4, and 11; see appendix Figure 52 to Figure 55 for all participants). The overall shape of the curves clearly reflects the recognition of "self" at morph-levels with large amounts of "self" (left sides of figures) and the recognition of "other" at morph-levels with small amounts of "self" (right sides of figures). If the test face was preceded by a prime that showed the face of the participant (prime self), the participants were more likely to press the associated "self"-key, which resulted in a rightward shift of the psychometric response function (solid lines). Inversely, if the prime showed one of the other faces, the participants were also more likely to recognize that other person, thereby shifting the curve to the left.

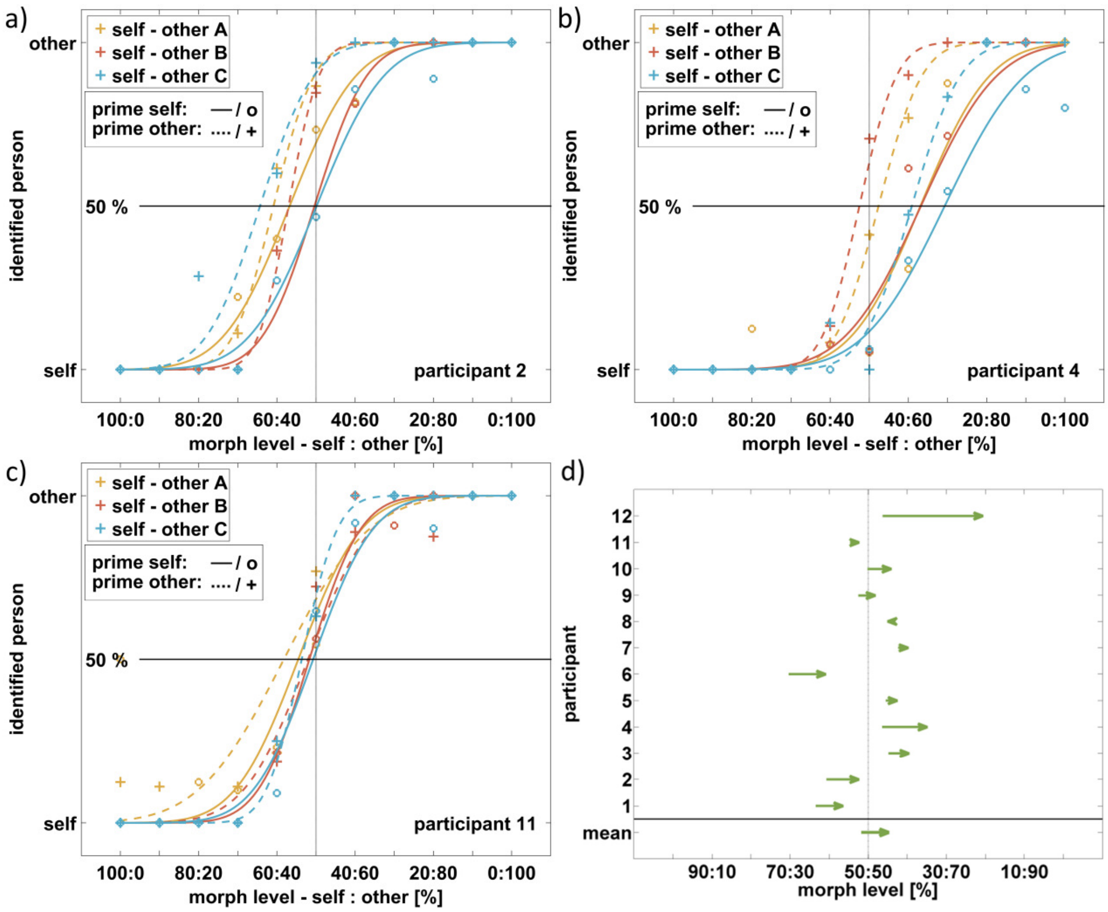

Figure 20. Identified person and induced PSE-shift. a) - c) Typical responses of the three participants (2, 4, and 11) for identifying the (test) person in dependence of the two different primes are displayed. Primes that showed the participant ("self", solid lines) resulted in more identifications as "self" than primes showing one of the other three persons (dotted lines) at a given morph-level. The markers (+ and o) show the average answers per prime and morph sequence at the given morph-level with psychometric curves fitted to these data. d) PSE-shifts for all participants are shown. Arrows indicate shift direction, magnitude and location where shifts occurred. On average, a significant shift of 7.1 % congruent with priming was induced. X-axes display the morph-levels in percent, y-axes display the identified persons or participants.

To determine the magnitude of this effect of altered response behavior, the PSE-levels (point of subjective equality) of the curves at both priming conditions were compared. For each participant, the mean of the PSE-values for the three curves with prime "other" and the three curves with prime "self" was calculated. These average PSE-values are shown as start- and endpoint of a vector in Figure 20 d); in other words, the beginning of an arrow is determined by the average PSE-level of the three curves with prime "other", and the endpoint of the arrow is determined by the average PSE-level of the three curves with prime "self". Except for one participant (number 8) all arrows point to the right side, which means the likelihood of identifying in the test picture the person who was presented as a prime was increased. On average, a significant shift of 7.1 % morph-level congruent with priming was evoked, $t(11) = 3.61$; $p < 0.01$.

In summary, both response times and face identification showed an effect of priming in the sense that the recognition of the primed person is facilitated and response time is reduced in congruent cases.

6.2. Experiment 3 b)

In the previous experiment 3 a), the effect of identity priming with self and other faces was investigated. Here in experiment 3 b), the effect of priming with other faces, which did not include the participants' faces, among each other was in focus. Furthermore, previous studies, e.g. by Bruce & Valentine (1986) and Ellis et al. (1987), could not find a priming effect of names. In contrast, Calder & Young (1996) did find such effects for names. This is why it was further investigated in this experiment 3 b) whether names can also induce a priming effect and if this is additional to the pictorial one of the displayed face-prime.

6.2.1. Material and methods

6.2.1.1. Participants

Twenty adults of European descent (nine male, eleven female) aged between 21 and 29 (mean age 23.3) with normal or corrected to normal vision participated in this experiment. They were paid a fixed amount of 8 € per hour. Again, care has been taken to select men without facial hair and women with no apparent facial make-up. Written informed consent procedures adhere to the guidelines of the Declaration of Helsinki and were obtained from all participants.

6.2.1.2. Stimuli and setup

The same setup and stimuli were used as in experiment 3 a). However, no personal pictures of the participants were taken this time since in experiment 3 b) the participants should judge between other people and not "self" vs. "other" like in experiment 1 or 3 a).

6.2.1.3. Procedure

Full instructions were given prior to the experiment. Before the main experiment started, the participants had to familiarize with the faces and names of three unknown persons and pass a learning test. Therefore, each participant was presented with unmorphed or "pure" 100 % images of three persons and was asked to assign names to them by typing on a keyboard. Participants were told to enter the true name in case they happened to know any of the persons. However, this case did not occur. Participants were told that they would be asked later in the experiment to recognize faces and names and that they should therefore study the faces and names carefully for as long as they wanted. The first part of a learning test began upon button press: One of the names was pseudo-randomly chosen and presented in the middle of the screen. The three faces were presented below and participants were asked to select the matching one by pressing the left or right arrow key to apply a cyclic permutation of the faces like "rotating a cylinder with pictures on it" and then space bar to confirm. Once each name was correctly assigned twelve times consecutively to the respective face, the second part of the learning test started. Again, the participants were presented with a pseudo-randomly chosen name followed by just one of the three faces. Then, they should answer 'yes' or 'no' via button-press (key 1 or 2 on keypad) if the presented face was matching the name or not. The learning test was passed if they succeeded six times in a row per face. This entire learning test lasted about five to ten minutes in total, depending on participants' performance and was carried out at the beginning of each session (see below). This was done to make sure that the participants familiarized themselves both with the faces and the names.

In the main experiment, each trial started with the presentation of one of the three names for 4 s pseudo-randomly selected, announced as target person (“Zielperson”). Here, the names were flanked by ‘X’s to the left and right so that the total number of letters (name + flankers) was kept constant with nine letters in all trials for equal visual input. Participants were instructed to “think of the person whose name was displayed”. Subsequently, a fixation cross was presented in the middle of the screen for 2-3 s (randomly chosen to avoid anticipation), followed by a prime image that was shown for 30 ms. Again, this duration corresponds to a periliminal stimulus. The prime image was always a 100 % picture. Then, the test picture from the morph sequence was presented for 170 ms. Afterwards, the screen went black. Participants were asked to determine if the shown picture (test) was the person whose name was presented at the beginning of the trial or not and press the according button on the keyboard in a yes-no-paradigm (Figure 21).

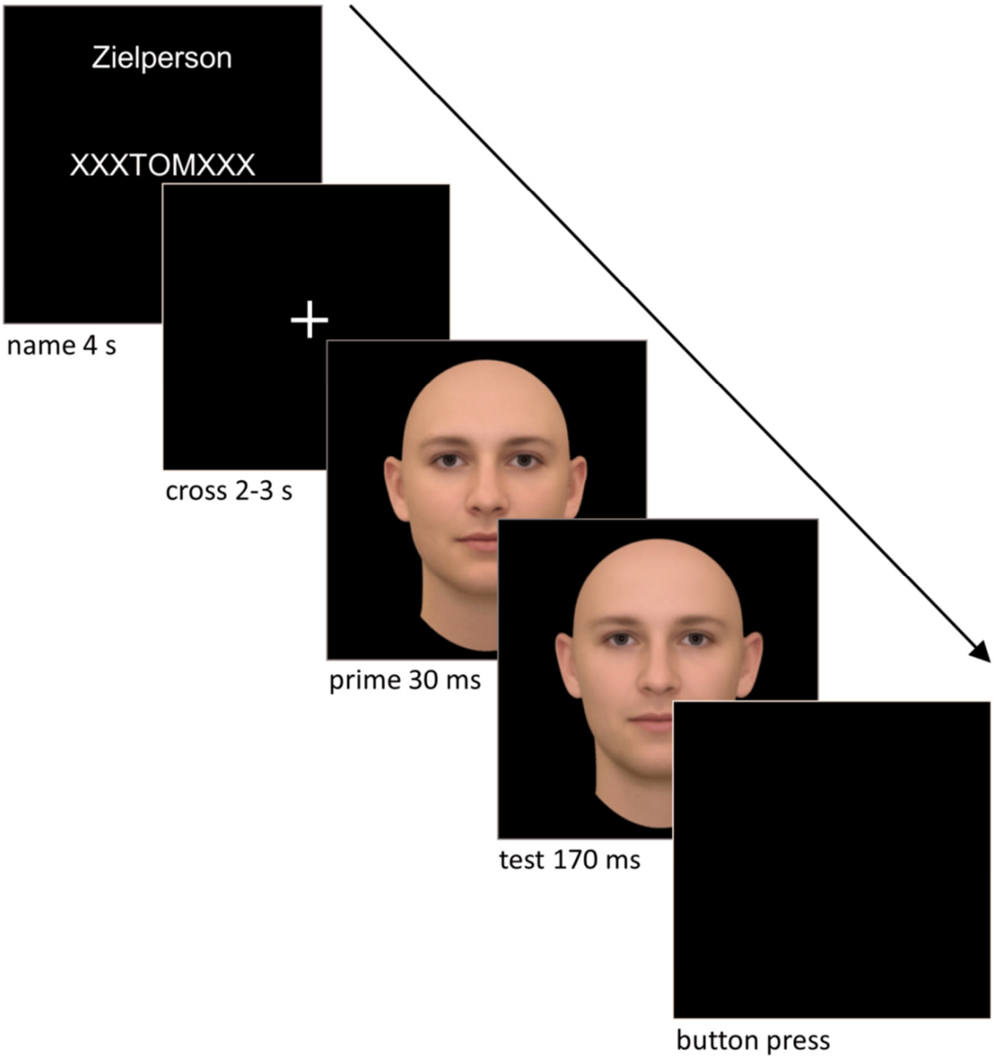

Figure 21. Sequence of a trial. Each trial started with the name of a target person who had to be imagined and identified in that trial. Then, a fixation cross was displayed for 2-3 s randomly chosen. Afterwards, the prime picture appeared for 30 ms. Then the test picture was displayed for 170 ms. After that the screen went black. The participant should indicate with a button press if the shown picture (test picture) was the target person whose name was displayed at the beginning or not.

Each prime - test combination was tested at least four times at each morph-level. Stimuli were presented in a randomized sequence. Additionally, a real-time evaluation of the responses was carried out to determine additional morph-levels for each combination for testing. The procedure to calculate the test levels was based on parameter estimation for sequential testing (“best PEST”, Pentland, 1980; Lieberman & Pentland, 1982) and is used here to gain more data around the inflection point of the psychological function. There were 420 trials in total for each participant to complete. These were split into two sessions lasting approximately one hour each and carried out on separate days.

6.2.2. Results

Trials with response times above 2.5 s were discarded (381 out of 8400). The limit was raised in comparison to experiment 3 a) since imagining a different person than oneself might take more time. No lower limit was introduced as the fastest response was 388 ms.

6.2.2.1. Response times

The average response time across all trials and participants was 1235.5 ± 423.4 ms. Like in experiment 3 a), fastest average responses occurred at low and high morph-levels (average minimum 1074 ms). Response times peaked around 50 % morph-level with an average maximum of 1415 ms. Generally, response times were larger than in experiment 3 a) (see Figure 19 and Figure 22).

In experiment 3 b) there were two possible target persons per morph pair that could be asked for, two priming pictures and eleven test pictures (morph-levels). In Figure 22 a) the response times for these combinations are displayed. The letters indicate the target person (first letter, A or B), the used prime picture (second letter, A or B) and the test picture (third letter, A to B in 10% steps).
In trials where target person, prime and test picture mentioned the same person [blue curve left side (A A A) and yellow curve right side (B B B) in Figure 22 a)] response times were smallest. On the contrary, if the test picture was different from the target person and prime [A A B, blue curve right side; B B A, yellow curve left side in Figure 22 a)] response times were largest. A mix of target person and prime picture regardless of test picture (A B x, B A x) led to intermediate response times [green and red curves in Figure 22 a)].

Here, the letters A and B stand for “one” or “the other” person during a trial. There is no “real difference” between the target person conditions A or B in terms of a fixed assignment of people to that letter, which is also reflected in the same response times and shapes of the yellow and red curves if they are mirrored left to right at the 50 % morph-level onto the blue and green curves, respectively. This was actually carried out with the data points and tested, resulting in no difference between them (Wilcoxon signed-rank test on the difference of the curves (blue - mirrored yellow, red - mirrored green) against zero, $p = 0.462$, $p = 0.795$). To investigate priming effects on the response times with respect to a target person in general, responses were therefore mirrored as described above and combined [red with green curves as well as blue with yellow curves; Figure 22 b)].

For the pooled data, again, a two-way repeated measures ANOVA of the factors prime and morph-level revealed a strong interaction, $F(10, 190) = 3.13$; $p < 0.001$, as well as a significant main effect of morph-level, $F(4.20, 79.85) = 19.56$; $p < 0.001$, but not of prime.

Greenhouse-Geisser corrected values were used for the morph-level as a violation of the assumption of sphericity was indicated for morph-level by Mauchly's test.

If the prime and test picture were showing the target person [e.g. A - A - A, blue curve left side, Figure 22 b)], response times were as low as 1074 ms. If target person and prime showed one person but the test picture showed another person [e.g. A - A - B, incongruent priming, blue curve right side, Figure 22 b)], response times were significantly increased to 1218 ms on average, $t_{100:0 - 0:100}(19) = 4.51$; $p < 0.001$. However, if the priming picture showed a different person than the target person (e.g. A - B - x), response times were equal on both lateral parts of the curve independent of the test picture (both sides of red curve).
Still, congruent priming on target person [blue curve left side, Figure 22 b)] led to faster and incongruent priming on target person (blue curve right side) to slower responses than any priming to non-target persons (red curve left and right side).

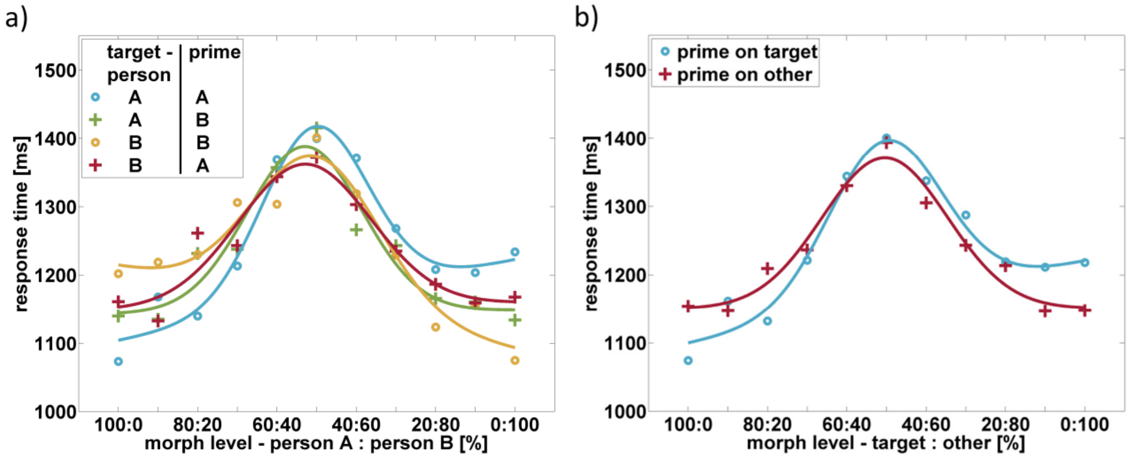

Figure 22. Average response times of all participants. a) For each "target person - prime" combination, separate average responses are displayed. Near the 50 % morph-level, it took participants longer to respond as the test image became more ambiguous. If target person and prime were different (green and red curve / cross data markers), response times at high and low morph-levels were not affected by primes as much as in cases in which target person and prime were the same (blue and yellow curve / circle data markers). b) Combined response times. Response times were smaller in cases where primes matched the target person and the displayed test picture (blue curve left side) and larger when non-matching (blue curve right side). No dichotomy of response times was present in cases where prime and target person did not match, irrespective of the test picture (red curve). Each marker (+ and o) shows the average response time per prime at that morph-level. The graphs were fitted combinations of a linear and Gauss distribution for each condition. X-axes display morph-levels in percent and y-axes response times in milliseconds.

6.2.2.2. Face recognition

Like in experiment 3 a), it was expected that priming would not only lead to modified response times, where congruent priming speeds up the responses, but would also manifest in a qualitatively different answering behavior. Thus it was investigated whether shifts in perceived identity were present in the participants' data. The mean shifts of the PSE values of all participants are displayed in Figure 23. Shifts were calculated like in experiment 3 a), that is, the difference between the curves at the 50 % level. Shifting direction again varied across participants. For four participants (12, 13, 14 and 18), arrows were pointing to the left side and therefore contrary to the hypothesis. Though, two of these four arrows were vanishingly small. For the other 16 participants, arrows were congruent with the primes and in accordance with the hypothesis. On average, a significant shift of 2.9 % morph-level in congruence with the respective priming was evoked, $t(19) = 2.88$; $p < 0.01$.

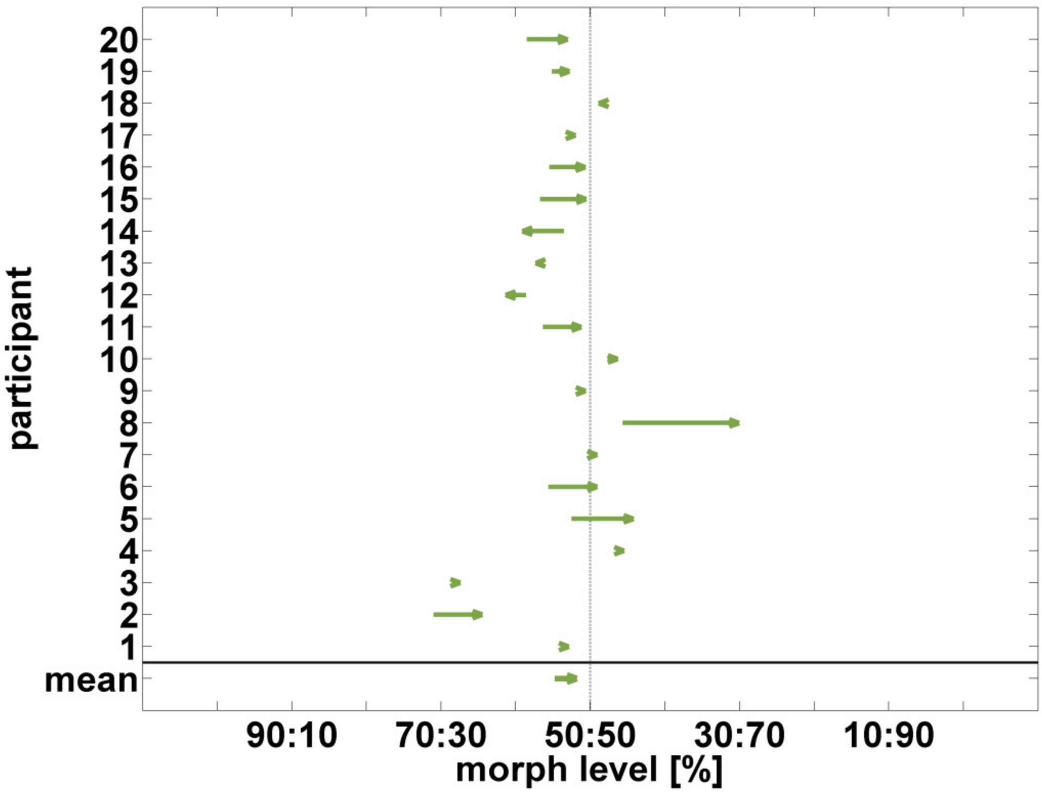

Figure 23. Induced PSE-shift. PSE-shifts for all participants are shown. Arrows indicate shift direction, magnitude and location where shifts occurred. On average a significant shift of 2.9 % was induced. X-axis displays the morph-level in percent, y-axis the participants.

Again, congruent priming on the target person led to faster responses and enhanced recognition of the primed person as response times and response behavior (PSE-shifts) revealed. However, priming of the non-target person did not result in faster or slower responses. Still, the increase of response time in the area with high ambiguity (around the 50 % level) was noticeable.

6.3. Discussion for experiments 3 a) & 3 b)

Two priming paradigms were presented in this chapter. In experiment 3 a), the participants were asked to make judgments about the identity of faces, namely, whether it was themselves or not. Response times and answering behavior displayed an effect of priming and confirmed the hypothesis. Responses were faster in cases with congruent prime - test stimuli. At subject-level, response times showed rather big variations (Figure 19); however, on average, bell-shaped response time curves with a congruency-dependent skewness were present. Here, two factors were expected to act upon the slopes: First, a factor of uncertainty that effects the bell-shaped curve since ambiguity is largest when two faces contribute to about the same amount in a morph picture. In contrast, at the ends/sides of the curve, it is quite clear which of the faces contributes the most and therefore is supposed to be recognized. Second, a factor of congruency between prime and test picture. If a prime and a test picture mainly show the same picture (and are combined with appropriate timings), a priming effect can be evoked that quickens the recognition of the second (here test) picture. According to Eimer & Schlaghecken (2003) and Pannese & Hirsch (2010), priming effects are dependent on different stimulus onset asynchronies and different prime durations, respectively. Both of these factors were present in the data and could be confirmed.

In the presented paradigm, priming was not only expected to have an influence on response times but also on the identification process in terms of "who" is recognized at a given morph-level. The effect of priming here led to a significant shift of the point of subjective equivalence (PSE), which means that depending on the prime picture participants identified (and therefore "saw") a different person.

According to the "interactive activation and competition" (IAC) model by Bruce and Valentine (see Figure 1), the presentation of a prime picture in the experiment would activate the "face-recognition unit" (FRU) of the respective person and eventually this person's "person identity node" (PIN), at least to some degree. Subsequently, the presented test picture - if it shows the same person - would also cause an activation of the same FRU. According to priming theory, the already pre-activated FRU would respond stronger/faster to the now-presented stimulus, and therefore also the connected and already pre-activated PIN would respond stronger. Taken together, this leads to a faster response of the participant for congruent prime - test pictures, which was also found in the experiments here.
In the incongruent case, a certain FRU (and PIN) would be activated upon priming. The test picture then would activate a different FRU and PIN. Activating a certain FRU and PIN also activates the other parts of the IAC network that are connected to this certain person while at the same time a suppression of the other non-matching connections takes place. Therefore, the (pre-) activation of the network for a certain person by the prime picture leads to a delayed activation once the test picture of another person is displayed, and consequently, the participant's response is delayed, too.
The "identified person" and the PSE-shifts of experiment 3 a) can also be explained by the IAC model. As explained above, the presentation of a prime picture activates the according FRU. The subsequent ambiguous test picture is supposed to activate both FRUs that are appropriate for the two persons the test picture consists of. Since one FRU was already pre-activated by the prime picture, it is more likely that this FRU and PIN reach the threshold in evidence accumulation earlier than the one responsible for the other person's face. Finally, the participant decides for the person who was also presented with the prime picture.
Similar effects were also visible in experiment 3 b) but with striking differences: Response times were generally larger than in experiment 3 a). This could be explained by the different paradigm of the experiment in which participants should think of a certain target person at first and then check whether the displayed person was matching. It is plausible that this imagining and verification process of other people is more demanding than just validating a picture to one's own image. Tong & Nakayama (1999) found a benefit of self faces in visual search, and also Ma & Han (2010) found evidence for faster self-face recognition as long as there is no self-concept threatening. However, a faster self-recognition in a direct comparison with "other" faces could not be seen [see experiment 3 a)].

Not only response times but also the magnitude of the PSE shift is lower compared to experiment 3 a), indicating a less pronounced effect of priming in terms of "who" the participants were recognizing.

The most pronounced difference, however, is the different courses of the response time curves. In experiment 3 a), the curves were symmetrical to each other. The amount of facilitation in congruent cases and inhibition in incongruent cases was identical. In experiment 3 b), response time curves were asymmetrical to each other. Congruent priming (here: congruency between identities) quickened the response only if the prime pictures showed the test picture and target person. In contrast, in cases where prime and test picture were the same but the target person who was asked for in the first place were different, the participants did not respond faster. This indicates that a second priming effect was taking place: Not only the prime picture affected the response on the test face but also the name of the target person seemed to influence the subsequent processing of (prime and test) pictures.
This result is remarkable since different studies found different results concerning priming effects of names on pictures: Bruce & Valentine (1986) found that such a priming effect is only present if the subsequent test also involves "naming" of the target person. If name retrieval was not necessary, no priming effect of names could be found. Also Ellis et al. (1987) stated that in familiarity judgments of faces, a priming effect was apparent if there was an earlier exposure to a picture of that person, but not if there was prior exposure to the person's name. However, Calder & Young (1996) found a cross-domain priming effect of names on faces, although, it was weaker than that of faces on faces (and names on names).
This seemingly inconsistent result can be explained by the IAC-model, too. First, the "target person" is shown to the participants as a name. This would activate the word recognition units (WRU), followed by the name recognition units (NRU) and the PIN of the matching person. Then, the prime picture is shown, leading to activation in the FRU and subsequently PIN. If there is a match, this person's PIN should be pre-activated. Finally, the test picture, which was asked to be identified, is presented. This again leads to activation in the FRU and PIN. If it is the same person in all three presentation cases, the decision output is faster. However, in cases where name and prime do not match, different PINs should be activated which also inhibit each other to some degree. If the test picture is presented then, there is no benefit of a faster decision.
In this case here, a small effect of matching prime and test pictures (and non-matching name) was expected since priming within domains (FRU - FRU) was expected to be stronger than across domains (NRU - FRU), but could not be detected.

How do the results of the experiments of this thesis match to previous findings of Bruce & Valentine (1986) in their second experiment and Ellis et al. (1987) who found no priming effects of names provided as text? Just as in Calder & Young (1996), Bruce & Valentine (1986) in their first experiment, and here in both experiments 3 a) and b), the names of the to-be identified persons were relevant to fulfill the task. If the names were unknown, participants would not be able to name the target person (Bruce & Valentine) or compare the identity of the target person as in Calder & Young and the presented experiments here.
In Ellis et al.'s (1987) and Bruce & Valentine's (1986) first study, however, participants were asked to make familiarity judgments. This does not necessarily require knowing the name of a person. We all know cases in which we remembered to have seen a certain person before (e.g. on TV) but we do not know the person's name, or that we do know the name of the person but simply cannot remember it right now (tip-of-the-tongue problem). Either way, the name is not necessary to perform familiarity judgment tasks.

The findings of experiment 3 b) can help to explain the different results of these studies: Names presented as text can cause identity priming effects on subsequent pictures - if those are relevant to solve the task. In the following, a unified model based on race-to-thresholds that can explain the results of these experiments will be presented.

6.3.1. Priming model

Reviewing the conclusions of the before mentioned studies and the results of the experiments presented here raises the questions when and how a priming effect of text gets established. In experiments in which naming was involved in solving the task [experiment 3 b); Bruce & Valentine, 1986; Calder & Young, 1996], an at least weak priming effect was present. In familiarity judgments (Bruce & Valentine, 1986; Ellis et al., 1987) in turn, no effect could be found. This indicates that a priming effect of written names is selective and depends on whether the name is relevant to solve the task.

In Figure 24, a model that visualizes the influence of the presented stimuli with regard to the different tasks is proposed. It consists of two instances, which could be networks or in the simplest case cells, that selectively respond on inputs and accumulate evidence; here, that would be the evidence for person A and person B. These would resemble the face-recognition units or a combination of FRUs and PINs of the IAC-model. The model presented here follows a basic race-to-threshold model where a certain threshold (green lines) needs to be reached for a decision to be made. Across time (x-axis), evidence (y-axis) for the two answering possibilities (person A, upper graph, or person B, lower graph) gets accumulated. Here, the urge to give an answer is included in diminishing thresholds over time. This ensures that an answer is given even if the presented stimuli are ambiguous. However, the confidence to have made the right decision is low then. Of course, this urge to answer could be modeled in a different way, too, e.g. with a constant threshold, but an (increasing) factor is added to the accumulated evidence with every time step so that it reaches the threshold faster (or reaches it at all).
The blue and red curves show the accumulated evidence for presentation of a person at a given point in time. Please note that blue means that the prime picture showed person A, and red means the prime showed person B - analogous to Figure 19 d) and Figure 22 b), and the accumulators for person A (“evidence A”) and person B (“evidence B”) are shown in the upper and lower part of the figure.

The three sections “instruction”, “prime” and “test” represent different stages of a trial. “Instruction” means the instruction of the task and the presentation of the target person as text. “Prime” is the period of time in which the prime picture was presented, “test” the display of the test picture.

In Figure 24 a), the expected time courses of experiment 3 a) are shown. The instruction (identify the person in the shown test picture as self or other) is not supposed to have an influence on the evidence accumulation. Upon presenting a prime picture (either A “self” or B “other”), the respective instance accumulates evidence while the other one does not. Then, the test picture is displayed. If it is congruent with the prime (blue curve in upper graph and red curve in lower graph) the evidence is already at an elevated level and therefore the threshold is reached soon (green dots at t_1) and the respective response is given by the participant. In incongruent cases, where the prime picture does not match the test picture, the evidence for a certain person, say A (blue lines), is also accumulated during the presentation of that prime picture (upper graph, blue line until end of section “prime”), but then, target person B is displayed and no further evidence for person A gets accumulated. In contrast, evidence for person B begins to rise (blue curve in lower graph, start of section "test"), but from a lower level, therefore more time is needed until the threshold is reached (green dots at t_2). These examples show the case for unambiguous test pictures (close to 100 % or "pure" images). If the morph-levels get closer to 50 %, the test pictures become more ambiguous, which should result in more shallowly rising curves in the test section and the thresholds be reached at a later point in time $t_n > t_2$ (not shown).

The model for experiment 3 b) is displayed in Figure 24 b). Please note that only the case in which the target name is A is shown. It is similar to Figure 24 a), but here, the instruction is supposed to have an influence already. Reading the name and thinking of the target person already seems to influence the evidence level (blue curve in upper graph). Therefore, the next stages of the trial (prime and test picture) start at an already elevated evidence level and then reach the threshold early (t_1) if congruent. In contrast, if the prime picture shows person B (red curves), the evidence for A stays at that level (dashed red line in upper graph) and eventually increases if the test picture is A again (red line in upper graph, t_2), or, if prime and test picture is B, they also rise and hit the threshold at t_2 (lower graph red line). In the last case, where target name and prime picture show person A, but the test picture is person B, the threshold is met last (lower graph blue line, t_3) because neither instruction nor prime would increment the evidence for B, resulting in flat lines until the test picture gets displayed.

These theoretical response times of the model match the recorded response times of both experiments 3 a) and 3 b), supporting its ideas. In figure Figure 24 c), the response time curves from Figure 22 b) of experiment 3 b) are shown again with added decision/response points t_1, t_2 and t_3 (green). The model fits well to the recorded data as it can explain the three different response time levels (blue left side = t_1 = fast, red both sides = t_2 = average, blue right side = t_3 = slow). For experiment 3 a) it is analogous [see Figure 24 a) and Figure 19 d)], however, only with two decision/response points t_1 and t_2.

For the model, the assumption was made that the rates of accumulation within each section are the same for both accumulation networks and only differ because of earlier events. Furthermore, other nonlinearities could apply to the curves, too. Also, please note that the scales of Figure 24 a) and b) may differ both between and within of them. Still, the outcomes and general inferences stay the same. Also, in typical race-to-threshold models, if no further evidence is provided (e.g. by presenting no stimulus), curves are thought to "decay" after a period of time ("leaky model", e.g. Usher & McClelland, 2001). Here, this would only be relevant to the section "prime" in experiment 3 b) [Figure 24 b)]. However, since the presentation time of the prime pictures was only 30 ms, the potential decay of this period of time should be neglectable.

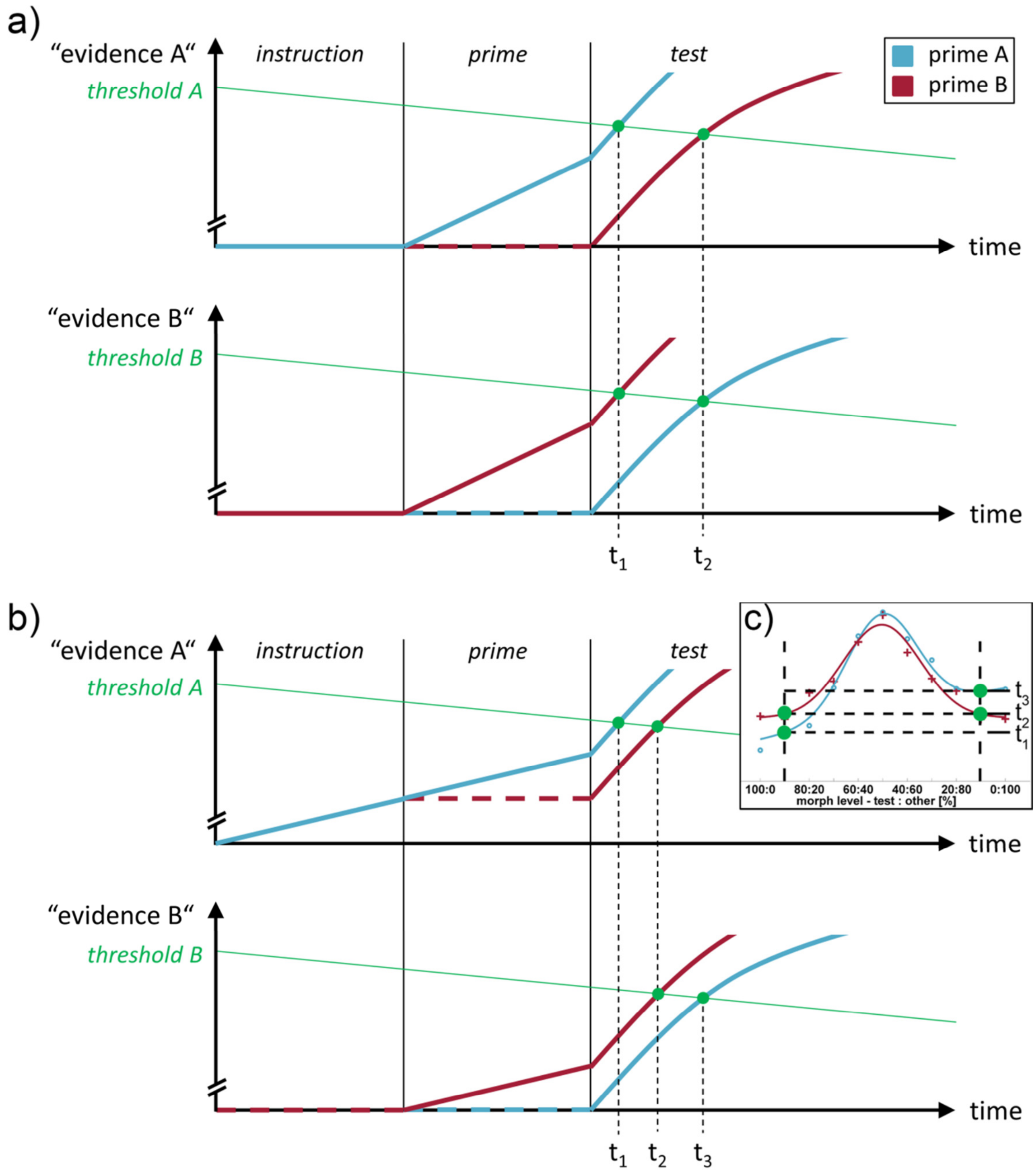

Figure 24. Priming model. The green lines display the thresholds that need to be reached for a decision to be made and an action to take place. The threshold is modeled to decrease over time to simulate an increasing urge to respond. Blue and red lines display the accumulated evidence when primed with a picture of person A (blue) or person B (red). a) Theoretical evidence accumulation for experiment 3 a) is shown. Depending on the prime and test pictures, evidence accumulation for person A or B meets the threshold first. b) Modeled time courses for experiment 3 b) are displayed. Only the case in which the target name is A is shown. Here, the “instruction” is supposed to impact evidence already. Dotted lines indicate non-existing extrinsic input for that accumulator or “resting evidence”. X-axes display time, y-axes display the evidence to decide for a given person (A or B). c) Average response time curves from experiment 3 b) and Figure 22 b) with added decision/response points t_1, t_2, t_3 (green). The model fits well to the recorded data and can explain the different time courses of the curves. X-axis displays morph-levels in percent and y-axis response times.

There are also other models that are used to explain decision making, e.g. the drift-diffusion model (DDM), which is similar to a random walk (Bogacz et al., 2006). In that model there is only one instance that accumulates different evidences together with (random) noise until one of several thresholds is reached. Depending on which threshold is reached, different decisions are made. However, results are expected to be comparable to the model with two (or more) racing accumulators as they all feature reaching a threshold by accumulating evidence. Furthermore, the model presented above fits better to the IAC model of Burton et al. (1993) as they suggest different units and nodes for different identities, while the DDM only features one instance.

One has to note that both models from Bogacz et al. and Burton et al. are typically used to describe two-alternative forced choice (2AFC) experiments, which might involve other mechanisms than yes-no-paradigms that were used here in this thesis (e.g. Bastin & Van der Linden, 2003). However, other studies say that this reported difference is caused by the type of experiment and different difficulties, and not because of any differences in brain mechanisms to actually fulfill these 2AFC- and yes-no-tasks - at least for patients with hippocampal lesions (e.g. Bayley et al., 2008).

In summary, the presented priming model can explain the results, and it fits to the data of the priming experiments on face representation presented in this thesis. It is in line with the IAC model of face recognition, and together they can also explain results of other studies in the field.

6.3.2. IAC model and discrimination task of experiment 1

In experiment 1, additionally to the categorization task, which was also used in experiment 3, discrimination tasks were carried out. Participants were shown three images of faces one after another, and they were asked to determine if the third face was the same as the first or second one. The question arises if the IAC model can be applied here, too, and what the processes would look like, or if the task was accomplished by mere image comparison.

It is possible to complete said discrimination task by mere image comparison, which means no person identification is necessary; instead, a pixel by pixel comparison between the pictures would be sufficient in the simplest way. All it needs is a working memory to store the presented images and compare them to the current input - which can be solved even by a simple computer program. However, such a computer algorithm would not show an effect upon crossing the PSE boundary like it has been found e.g. by Beale & Keil (1995), since the program does not know anything about “identity”. This in turn means that, at least for familiar faces, there is supposed to be more than a simple image comparison taking place. Although in experiment 1 no increased discrimination performance could be found when the PSE level was crossed, account will be given to the activation patterns in the IAC model if it responds to the stimuli:

The IAC modules FRUs and PINs should be activated according to the stimulus strength they receive - which means with an e.g. 30:70 % morph picture the respective FRUs (and PINs) should be activated by 30 % and 70 % in theory. However, since there are also inhibitory connections, they are supposed to be interacting and therefore altering each other’s activation level - probably with an advantage for the stronger activated part (here 70 %). Once the next image gets displayed (e.g. 50:50 %) the same FRUs and PINs receive input, though, with different strengths than before. It is feasible that the FRU that previously had the stronger input still has a benefit here as the activation before was stronger. However, it could also be that the change of stimulus strength benefits the other FRU as there was an increase in stimulus strength from the first to the second picture in that one. Finally, the third image gets displayed, which again stimulates both FRUs and

PINs, and it is mentally compared to the previously presented images. Depending on the morph-level of this picture and the potential pre-activations, a response is made.
With the above mentioned considerations, differences in the participants' responses could occur as well as further unknown activations and interactions may take place, which also could influence each other.

To investigate this, participants' response times were sorted as follows: Only image pairs that did not touch the 50 % boundary were analyzed with respect to response time and answering behavior (correct and false response), which means image pairs 30:70 % - 50:50 %, 40:60 % - 60:40 % and 50:50 % - 70:30 % were excluded since they were at or going across the boundary.
Furthermore, image pairs were divided into two groups: In the first group (named "pure") there were images in which the third and therefore to-be matched image was closer to the pure 100 % images, e.g. 10:90 % - 30:70 % - <u>10:90 %</u>, or 60:40 % - 80:20 % - <u>80:20 %</u>. In the second group (named "ambiguous") there were images in which the third image was closer to the ambiguous 50 % level, e.g. 20:80 % - 40:60 % - <u>40:60 %</u>, or 70:30 % - 90:10 % - <u>70:30 %</u>. Results for correct and false responses are shown in Figure 25.

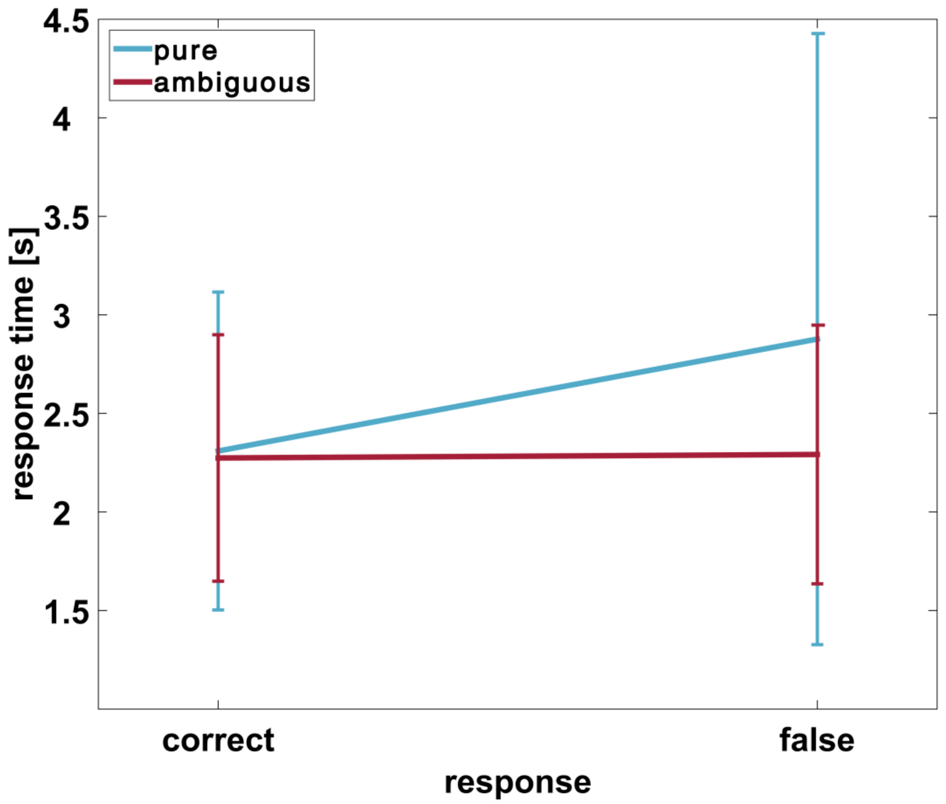

Figure 25. Effect of ambiguity on responses. Data were sorted into two groups in which the third picture of the sequence was either more "ambiguous" (red) or more "pure" (blue). Response times were at the same level unless the third picture was more "pure" and participants were answering falsely. In this case, there was a tendency for increased response times. X-axis displays participants' responses and y-axis response times in seconds. Error bars show standard deviation.

If participants matched the third person correctly, their response was counted as “correct”, if not it was “false”. In cases where the third picture was the more “ambiguous” one (red), response times were at the same level for both correct and false answers. If the third picture was the more “pure” one (blue) in the image sequence and the answer was correct, response times were at that level, too. However, if the third picture was more “pure” and the participants were answering falsely, there was a tendency for increased response times as a two-way repeated measures ANOVA of response and group revealed in a weak interaction, $F(1, 9) = 3.926$; $p = 0.079$.

What does this mean for the IAC or image comparison strategy? Those cases in which there was a tendentially larger response time (blue, false), they had image sequences that consisted of more “pure” third images and a false response was given, e.g. 40:60 % - 20:80 % - 20:80 %, or 100:0 % - 80:20 % - 100:0 %. It is just a tendency in the data, but that cannot be explained by a simple image comparison strategy as it does not account for such a biased answering behavior. Such a strategy would not distinguish between false answers to one or the other presented picture. With more “pure” images, there was a tendency that it took participants longer to respond falsely. This indicates that recognition processes are involved since participants seemed to be (subconsciously) hesitating to push the (false) key. Probably a (false) decision was already made but a recognition/identity module was interfering with the execution of the action. Still, the decision to push that certain key was already made and then the action was carried out - but with a delay. This result supports the idea of a recognition based mechanism to solve the discrimination task and that it is not a simple image comparison, e.g. like computers would carry out.

The experiments presented so far in this thesis highlight commonalities and differences between direct identity priming of self and other faces, the impact of “self” in recognition tasks, and the task dependence of stimuli with regard to their priming characteristics. Also, the face-space and its characteristics as well as naturally occurring shifts were described. The priming model illustrates time courses and gives an account of the response times. These findings are in line with and support the IAC model of face recognition.

This closes the first part of this thesis. In the second part, the representation of places will be addressed. In the general discussion chapter of this thesis, the insights on face recognition and representation will be revisited.

6. Experiment 3 - Primed face recognition

7. Introduction for experiments 4* and 5

In this part of the thesis, it is investigated how human navigators orient themselves and how the current location of "self" influences the spatial representation of locations recalled from long-term memory. In comparison to the representation of faces, an account on the representation of places will be given along with navigation strategies and models as well as involved brain structures.

Walking and navigating in different environments and spaces are crucial skills for humans. As we walk through an environment, we constantly keep track of objects, landmarks and path opportunities around us. This environmental information forms a working memory (for working memory see Baddeley & Hitch, 1974; Baddeley, 2000) of surrounding space for which Loomis et al. (2013) suggested the term "spatial image". In this spatial image, positional information about objects in the surrounding environment is stored in an ego-centric representation. Ego-centric representation means that positional information of objects is stored relative to the observer who is in the center of this representation. Objects then are represented to be "behind" or "2 meters ahead" or "45 degrees to the left" etc. Local, ego-centric representations of space have been studied in many contexts, including among others sensorimotor integration, visual scene recognition and spatial cognition. Tatler & Land (2011) and Land (2014) reviewed a large body of evidence on ego-centric visual representation supporting the stability of perception across eye-movements as well as eye-hand coordination with and without locomotion of the body. The representation considered by Tatler & Land (2011) extends around the agent up to about the size of a room in an indoor environment. A similar spatial working memory including also a mechanism for spatial updating has been suggested by Byrne et al. (2007). This spatial updating takes place automatically as one turns and moves around and cannot be suppressed nor be elicited by imagined movements (Farrell & Robertson, 1998). The notion of the spatial image (Loomis et al., 2013) is slightly more general in that it may include knowledge from other (non-visual) modalities and extends to more distant spaces, which may be out of sight even if the observer would turn his or her head accordingly. Information from distant locations beyond the current sensory horizon can originate from two sources, i.e. long-term memory of distant places, or spatial updating if the distant place had been visited before and was since maintained in working memory. Such a long-term memory of distant places and larger environments is often referred to as a "cognitive map" (Tolman, 1948; O'Keefe & Nadel, 1978).

---

* This chapter and experiment 4 have been published as an original research paper in Röhrich et al. (2014). The publication is available at *journals.plos.org/plosone/article?id=10.1371/journal.pone.0112793*. Minor changes were applied to the chapter for this thesis. The raw data were collected by Niklas Binder ("Holzmarkt") and Julia Mayer ("Marktplatz") who also used the data for their bachelor thesis and state exam essay, respectively. Both the programs for analysis of the data were programmed and the analyses were carried out by me. No parts of their works have been used in this doctoral thesis.

7.1.1. Multiple representations of space

Multiple representations of space have been suggested for a number of reasons. One issue is the problem of scale, which may vary from centimeters in manipulation tasks to thousands of kilometers in way-finding. Grusser (1982) distinguishes a (mostly metrical) grasp space, a near- and a far-distance action space, and a visual background. Montello (1993) presented a classification of "psychological spaces" also based on scale in which the spatial image discussed here is somewhere between "vista space" (what is currently visible) and "environmental space" (the area a subject is used to navigate in).

The distinction between working and long-term memories of space is grounded both in behavioral and neurophysiological data (Carruthers, 2013; Chen et al., 2013). Spatial working memory tasks which are largely independent of spatial long-term memories include spatial sequence learning such as walking versions of the Corsi block-tapping task (Piccardi et al., 2013; Röser et al., 2016), perspective taking and spatial updating (Farrell & Robertson, 1998), walking without vision (Philbeck & Loomis, 1997), path integration (Loomis et al., 1993; Wolbers et al., 2007), path-planning in multi-local tasks (Hardiess et al., 2011), etc. Interactions of spatial working and long-term memories are crucial in wayfinding, i.e. the planning of novel paths from known segments (Hartley et al., 2003; Meilinger et al., 2008; Wiener & Mallot, 2003), spatial imagery (Mellet et al., 2002), direction giving and other tasks. Wang & Brockmole (2003) studied spatial updating, a typical working memory task, in nested environments and concluded that spatial updating acts differently on close (the surrounding room) and distant (the outdoor buildings) environments. Giudice et al. (2013) addressed the interaction of long-term and working memories in a pointing task involving the angle between items stored in the different memory systems.

In a study by Basten et al. (2012), visitors of the University restaurant of the University of Tübingen were asked to draw sketches of the timber market ("Holzmarkt"), a central and familiar downtown square about two kilometers away. Drawings were rated for orientation and a clear preference for the southward view was found, depicting a landmark church building on top of a hill. However, when participants had been asked prior to the sketching task to imagine walking a route passing by the target square in one of two opposite directions, drawings in the respective viewing direction became significantly more frequent. The authors concluded that mental travel activated a view-dependent ("ego"-centric with respect to the imagined travel) representation of the target square which later primed the sketching process.

A particularly interesting case for the present discussion is representational neglect (Bisiach & Luzzatti, 1978), which shows that (at least in patients suffering from hemilateral neglect) recall of spatial long-term memories depends on the subject's imagined position and orientation. One obvious interpretation of this finding is that recall from long-term memory goes into some sort of spatial image or working memory centered at the observer's imagined position and that it is the left side of this representation which is affected by neglect.

Spatial memory systems may differ in the reference system employed to organize spatial information. Perception is ego-centric and so is the assumed spatial image (Burgess, 2006; Loomis et al., 2013; Tatler & Land, 2011). In perspective taking, route planning and mental travel ego-centric memories centered at imagined positions may also exist. The reciprocal term, allocentric, is harder to define. Summarizing discussions e.g. by Klatzky (1998), Burgess (2006) and Mallot &

Basten (2009), here in this thesis, an allocentric memory is defined as one that does not change as the observer moves. Note that this definition does not refer to coordinate systems or global anchor points. Indeed, knowledge such as distances between places as well as oriented views and their relation to other oriented views qualifies as allocentric memory in this sense, because it can be carried around and remains useful without a need for movement-dependent changes or transformations. Almost as a corollary to this definition, long-term memories will always be allocentric, while working memories involving automated spatial updating will be not. In the model section, the view-graph (Schölkopf & Mallot, 1995) is described as an allocentric data structure for spatial long-term memory that lends itself easily to interactions with ego-centric working memories.

Over the past decade, imaging studies have identified an extensive network of cortical and subcortical brain areas involved in a variety of spatial behaviors. Tasks involving an interplay of spatial long-term and working memories have been shown to recruit structures such as the retrosplenial cortex as well as medial temporal lobe (Bird & Burgess, 2008; Ranganath & Ritchey, 2012; Vann et al., 2009; Wolbers & Buchel, 2005). More on the visual side, scene recognition as well as imagery of out-of-sight places or perspectives has been related to various parts of the parietal cortex and transverse occipital sulcus (Lambrey et al., 2012; Nasr et al., 2013; Schindler & Bartels, 2013).

7.1.2. A view-based model of spatial working and long-term memories

In the interplay between spatial working and long-term memories, the encoding, or data-format, used by each memory structure is of great importance. Recall from long-term memory into spatial working memory, i.e. between allocentric and ego-centric representations, is often thought to require a coordinate transform or reference frame transformation. In reference frame transformation, a spatial layout with a certain orientation gets transferred into a different layout with a different orientation while all the objects inside the layout keep their relations between each other. An example would be the transfer of spatial knowledge from a printed map (bird's-eye view) into walking directions in the environment (first-person view); for an overview see Klatzky & Wu (2008). This transfer is certainly true if spatial information is explicitly represented in the form of coordinates. However, in a view-based account, an allocentric, long-term representation of place may even be a view or a collection of views which were egocentric when first perceived and stored, but are now carried around for reference. Simply enough, transformation of this view-based allocentric representation into an egocentric one amounts to picking a particular view which corresponds to the current viewing direction and loading this view into working memory, e.g. for comparison to the currently visible view of the present place. As a result, places would be recognized by view matching (Gillner et al., 2008), similar to the snapshot algorithms discussed in insects (Cartwright & Collett, 1982). In addition to simple matching, a process of view-transformation might be involved, allowing the prediction of nearby or intermediate views from stored ones, as has been suggested for robot applications (Moller et al., 2010). Such a mechanism seems to be required also in the pointing task studied by Giudice et al. (2013), involving both long-term and working memories. In pose-invariant object recognition, view interpolation is a well-established mechanism (Bulthoff et al., 1995; Ullman & Basri, 1991).

The concept of view-based representations of navigational space has been developed by Schölkopf & Mallot (1995) and used in robot simulations (Franz et al., 1998) and models of hippocampal processing (Gaussier et al., 2002). Behavioral evidence for view-based navigation in humans has been presented by Mallot & Gillner (2000), Wang & Spelke (2002), and Pickup et al. (2013). View specific neuronal activity has been reported e.g. from the monkey parahippocampal formation (Furuya et al., 2014) or the human retrosplenial cortex (Wolbers & Buchel, 2005).

7.1.3. Model

The central spatial concept of the view-based framework is the view, i.e. an image or early visual representation of a sector or angle of the environment taken at a position $\boldsymbol{x} = (x_1, x_2)$ and with a viewing direction φ; the view is denoted by $v(\boldsymbol{x}, \varphi)$. It needs not be limited by the visual perimeter, but may also contain information from beyond the current visual horizon, encoded in an egocentric way, see, for example, Tatler & Land (2011). The simplest long-term memory of a place $\boldsymbol{x}_o$ is then a collection of views taken at that place, $\{v(\boldsymbol{x}_o, \varphi_i), i = 1, \ldots, n\}$, where the index i enumerates the individual viewing directions and n is the total number of views stored for the particular place [see Figure 26 a)]. The views may be overlapping and the distribution of viewing directions φ_i may be anisotropic. If, for example, one particular view of a place is especially salient, this may be modeled by assuming that multiple copies of this view, or largely overlapping adjacent views, will be included in the place representation. In analogy to object representation, such views might be called "canonical" for the respective place. In addition to the views themselves, here, it is assumed that the adjacencies of views are also represented in the place code. The views together with their adjacencies thus form a simple view-graph with a ring-topology. As in Schölkopf & Mallot (1995), the adjacency links will be labeled with action codes such as "turn left", or "turn right 40 degrees".
From this place representation, a long-term memory of a larger environment, i.e. a cognitive map, can be built as a full view-graph and used for way-finding [see Figure 26 b)]. For multiple places, inter-place view adjacencies have to be stored as "action labels", representing egocentric locomotor actions such as "walk straight from here" or "follow the street from here". In these action labels, "here" refers to a view from the current place assuming the observer's current heading. The link will end at a view of a neighboring place, as it appears when arriving from the starting location. As was demonstrated by Schölkopf & Mallot (1995), the resulting view-graph contains sufficient information for route planning and navigation between connected views by means of the associated actions.

As a model of spatial working memory, a sub-graph of the full view-graph is suggested, consisting of the current view corresponding to the observer's current position and orientation, and the views reachable from this current view in a small number of steps s, i.e. the outward neighborhood $N_s(v_o)$. Note that the graph links are directed, allowing to distinguish an outward neighborhood (views reachable from v_o) from an inward neighborhood (views from which v_o can be reached). In Figure 26 c), the one-step ($s = 1$) outward neighborhood of view 1 of place B is shown. As the observer moves, the current view will change and so will its outward neighborhood represented in working memory.

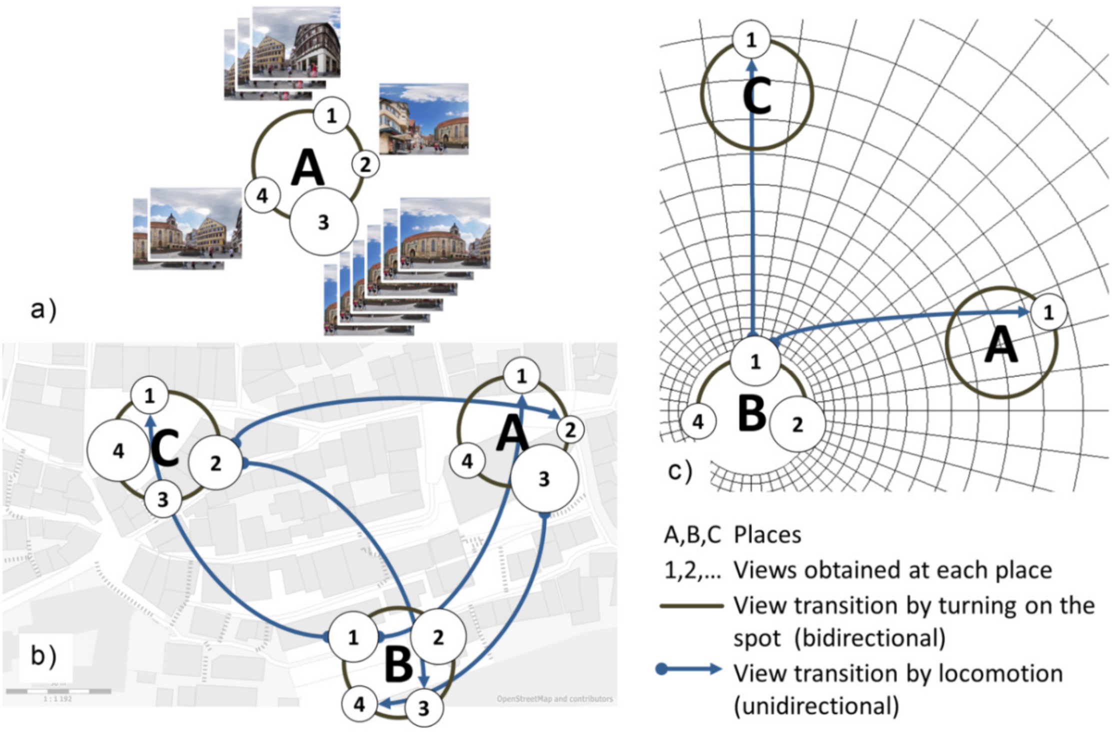

Figure 26. View-based model of spatial long-term memory. Upper case letters A, B, C denote places, numbers (1 - 4) denote views visible at each place. E.g., view A3 depicts a church building when standing at the "Holzmarkt" (A), facing south. a) Place representation composed of a collection of directional views (1-4) obtained at a place A. Views may be represented multiply or overlapping, allowing to represent viewing direction in a population code. The size of the circles indicates the frequency with which each view is stored or the likelihood that it is read out in recall. (Tübingen "Holzmarkt" icons are sections of a panoramic image retrieved with permission from www.kubische-panoramen.de.) b) View-graph of twelve views (A1 – C4) belonging to three places. Within each place, views are linked by turning movements. Views of different places are linked by movements involving translations. Note that these links are unidirectional; for example, a path from A to B starts from view A3, while the return from B to A will end on A1. c) A view-based model of spatial working memory is obtained by extracting a sub-graph from the total view-graph. It contains the current view (B1) which also marks the current observer position and forward direction and its outward neighborhood of order 1, i.e. the directly adjacent views (A1, B2, B4, C1). Outward neighborhoods of higher orders may also be represented but are not shown in the figure. The polar grid is added to indicate that metric updating may take place in the working memory, which, however, does not play a role in the experiment reported in this thesis. Map source: © OpenStreetMap contributors.

This may be achieved by repeatedly refreshing the neighborhood from long-term memory, i.e. loading the appropriate sub-graph after each movement step. Alternatively, or on smaller scales, one could think of some sort of ego-motion driven image transformation (spatial updating) within working memory. Here, this possibility is indicated by adding a polar coordinate grid to working memory in Figure 26 c). In the experiments, one cannot distinguish between refreshing from long-term memory and spatial updating within working memory. See Giudice et al. (2013) for an experiment directly addressing this problem.

7. Introduction for experiments 4 and 5*

When asked to imagine a nearby target place $\boldsymbol{x}_t$, participants will recall from memory one of the stored views $v(\boldsymbol{x}_t, \varphi_j)$ of this place. In spatial working memory, only the views contained in the outward neighborhood of v_o will be present. Therefore, if recall is based on working memory content, the view obtained when (mentally) traveling from the current "here" to the target place will be selected. In this case, it is predicted that in visual recall of a target place, the recalled viewing direction will depend on interview location. If, however, recall is based solely on long-term memory, one of the known views of the target place will be selected independent of interview location.

For the analysis of the data presented below, the following notation is introduced: Let $p_{i,t}(\varphi)$ denote the probability that the recalled view of target place $\boldsymbol{x}_t$ has the orientation φ, given that the interview location is $\boldsymbol{x}_i$. Let further $L_t(\varphi)$ and $W_{i,t}(\varphi)$ denote the probability densities of recalling a view φ if recall is from long-term and working memory, respectively. Note that the working memory contribution depends on interview location, whereas the long-term memory contribution does not. $W_{i,t}(\varphi)$ is expected to be a peaked distribution with a maximum at the approach direction from interview location $\boldsymbol{x}_i$ to target place $\boldsymbol{x}_t$. In the data analysis, the approach direction will be identified with the air-line direction between the two places,

$$\varphi_{i,t} = \operatorname{atan2}(\boldsymbol{x}_t - \boldsymbol{x}_i) \tag{3}$$

where atan2 is the inverse tangent function with two arguments. For the distribution of the recalled view orientations, we obtain

$$p_{i,t}(\varphi) = \alpha\, L_t(\varphi) + (1-\alpha)\, W_{i,t}(\varphi) \tag{4}$$

where $L_t(\varphi)$ and $W_{i,t}(\varphi)$ are the long-term and working memory contributions, respectively, and α is a mixing factor varying between 0 and 1. It reflects the relative strength of long-term and working memory components in the recall. α is expected to be less than 1 for interview locations close to the target place and 1 for distant interview locations.

If, for a given target place, the interview locations are spaced regularly around this place, the average of the $W_{i,t}(\varphi)$ will approach the uniform distribution, $(1/n)\sum_{i=1}^{n} W_{i,t}(\varphi) \approx 1/2\pi$ and the long-term memory contributions are estimated as

$$\alpha\, L_t(\varphi) \approx \bar{p}_{\cdot,t}(\varphi) - \frac{1-\alpha}{2\pi} \tag{5}$$

where $\bar{p}_{\cdot,t}(\varphi)$ denotes the average view distribution over all interview locations. From this, an estimate for the working memory contribution will be calculated as

$$W_{i,t}(\varphi) \propto p_{i,t}(\varphi) - \bar{p}_{\cdot,t}(\varphi) + c \tag{6}$$

where c is a constant reflecting the non-zero average of the working memory distributions. In the analysis of the experimental data, orientations are sampled to the four cardinal directions (N, E, S, W). The constant c cancels out in the calculation of the circular vectors following Eq. 7 below. In analyses of the distribution $W_{i,t}(\varphi)$, this constant is important to avoid negative values. Therefore, it can be set to 0.25. The proportionality factor in Eq. 6 will be ignored in the sequel.

8. Experiment 4 - View-based spatial memory

Experiment 4 consists of two experiments. In both of them it was investigated how participants innately respond if they are approached and asked for a quick sketch of a known place while they are located in the area. For experiment 4 a), this known place was the timber market ("Holzmarkt"), and for experiment 4 b), it was the market place ("Marktplatz"). Both squares are well-known places in the city center of Tübingen (Germany). The data was evaluated with respect to the different interview locations around town and compared with the model suggested above.

8.1. Experiment 4 a) - "Holzmarkt"

For orientation and navigation, it is not sufficient just to know where other things are, but it is crucial to know where oneself is currently located. While navigating in a familiar environment, local knowledge from allocentric long-term memory is transferred to working memory involving reference frame transformations. One way to retrieve information from participants about stored long-term knowledge of familiar places and have it transferred into working memory is to ask them for directions. However, this approach has drawbacks, e.g. the limitations of verbalization of such knowledge and also the assumptions the asked person makes about the local knowledge of the person who is asking. This will lead to simplified answers that are indeed helpful for people who actually need navigational advice, but it does not tell much about the content of the participants' memory systems. Here in experiment 4, participants were asked to draw. This method removes language limitations and allows them to spend more time contemplating. Further, they were not asked to give directions, which removes the urge to simplify the answer so that the other person can understand or keep the advice easier.

It is investigated if the drawn sketches have a general orientation or if they vary. It is hypothesized that the sketches' orientation is dependent on the interview location, indicating an interaction between one's current position and recall from long-term memory.

8.1.1. Material and methods

8.1.1.1. Participants

Passers-by at 14 locations in Tübingen (see below and Figure 27) were approached during day time and asked "if they would participate in a quick interview for a navigational study". They were informed about the type of the collected data and the general procedure. About one third agreed to participate (verbal informed consent) as was documented by their later participation in the interview. Participants were not asked for their names and accordingly were not required to give their consent in writing. Participants were free to terminate their participation at any time, simply by walking away. The informed consent procedures adhere to the guidelines of the Declaration of Helsinki, approval by the local ethics committee was not required.

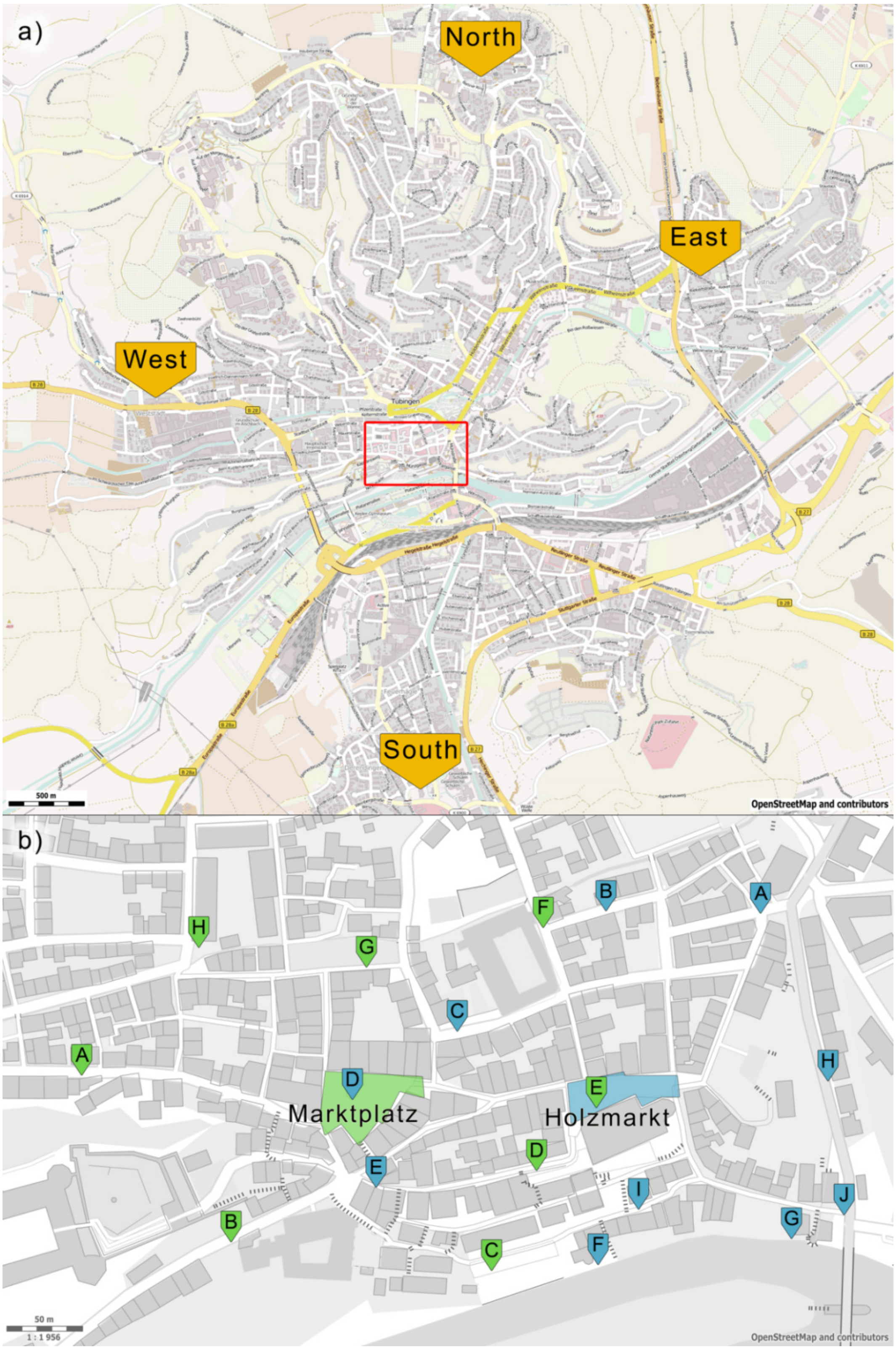

Figure 27. City maps of Tübingen with interview locations and target places ("Holzmarkt" & "Marktplatz"). a) Distant (suburban) interview locations (north, east, south, west) were located in small shopping areas about 2 km away from the target squares, which were inside the downtown area (red square). b) Close-up view of the downtown area of Tübingen. Blue: Interview locations (A - J) and target place for experiment 4 a) ("Holzmarkt"). Green: Interview locations (A - H) and target place for experiment 4 b) ("Marktplatz"). Map source: © OpenStreetMap contributors.

8.1.1.2. Procedure

Participants were requested to “sketch the layout of the timber market” (“Holzmarkt”), a well-known down-town square, on an A4 sheet of paper. After sketching, they were asked for their age, years of residency in Tübingen, own judgment of general navigation skills and own judgment of local knowledge (see below). Only sketches by participants who had lived in Tübingen for more than two years were analyzed further. In total, these were 335 adults (161 male, 174 female). An interview and sketch map production took less than two minutes in total. Examples of sketch maps are shown in Figure 28.

Interviews took place outdoors, either at one of four distant locations in small suburban shopping areas about 2 km away from the target square (“distant” condition) or at one of ten downtown locations in walking distance (about 150 m) to but out of sight of the target square “Holzmarkt” (“near” condition; Figure 27). Also see Figure 56 for an aerial photograph of the downtown area with the target place and nearby interview locations. Care was taken to approach participants walking in different directions. Approach was from sideways with respect to the participant’s heading. Upon being approached, participants stopped but did not change their general body orientation. Also during recall, no regular turning movements of the participants were observed. Body turning could influence the results both by physically aligning with the target place and priming that orientation, and on an internal processing level on mental rotation (Lohmann et al., 2017) and therefore was not desirable.

The sketches were categorized for orientation (north, east, south or west up) by three independent raters. From the 335 drawings, 331 were judged identically (99 %) with a chance-corrected inter-rater reliability of $\kappa = 0.98$. Only the 331 identically judged drawings were analyzed further (254 near condition, 77 distant condition). The mean age of the 331 participants whose maps were included was 33.36 years, their average time of residency in Tübingen was 12.9 years, their own judgment of local knowledge and general navigation skills was 5.9 and 6.2, respectively, both on a scale between 1 and 9 with 1 = very poor and 9 = very good.

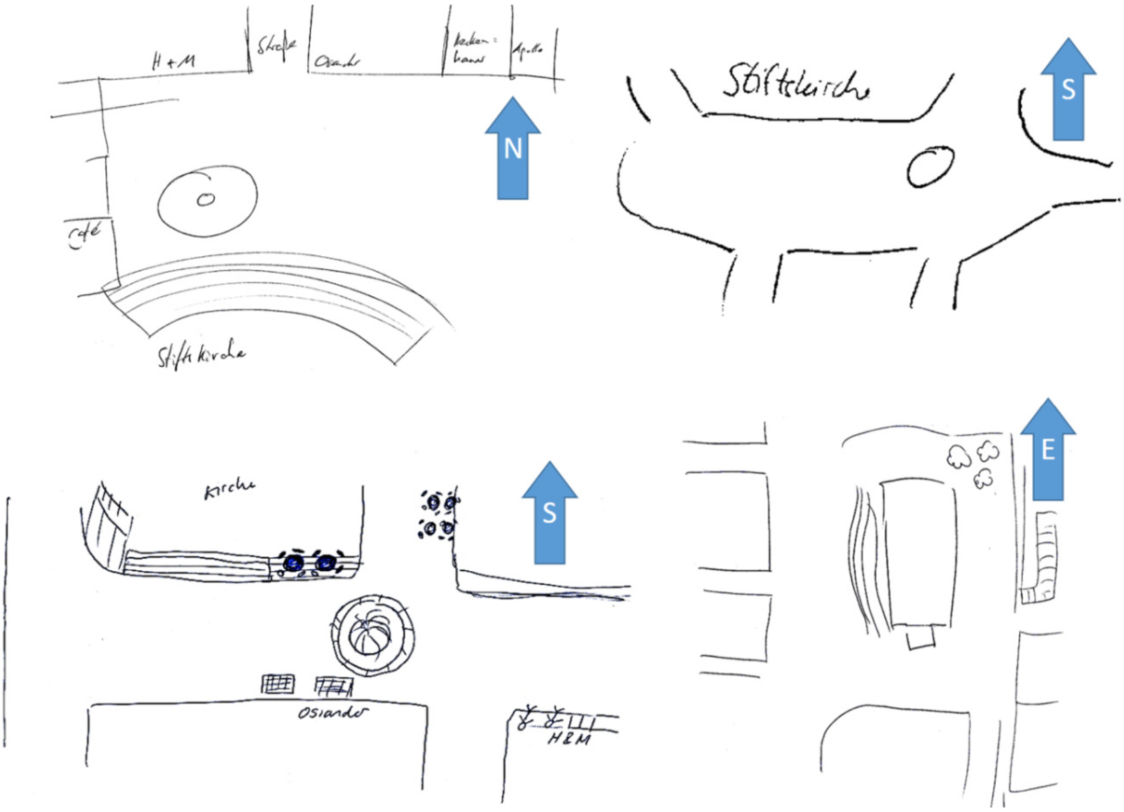

Figure 28. Examples of sketches of the "Holzmarkt" from four participants. The blue arrows indicate the orientation the sketches were rated in by three raters during analysis. Note the inscriptions "Stiftskirche" or "Kirche", referring to the landmark church building located at this square [see also view A3 in Figure 26 a)]. The "parallel" lines mark a flight of stairs leading from the square to the church, the circles mark a fountain at the Western side of the square.

For each interview location i, relative frequencies of ratings for the four cardinal directions were calculated and denoted as (n_i, e_i, s_i, w_i) for north, east, south and west. Average frequencies were also calculated separately of the ten "near" and the four "distant" interview locations $(\bar{n}, \bar{e}, \bar{s}, \bar{w})$. In the next step, the average frequencies from the "near" interview locations were subtracted from each of the local histograms of the "near" condition [blue in Figure 27 b)]. Similarly, the average frequencies for the four distant interview locations were subtracted from the distant histograms [yellow in Figure 27 a)]. The results are referred to as the "location-dependent components" and considered as an estimator of local working memory content, according to Eq. 6. Finally, these location-dependent components were transformed into location-dependent orientation vectors

$$w_i = \begin{pmatrix} (e_i - \bar{e}) - (w_i - \bar{w}) \\ (n_i - \bar{n}) - (s_i - \bar{s}) \end{pmatrix} \tag{7}$$

The orientation of these vectors is an estimator of the circular mean of the working memory distribution $W_{i,t}(\varphi)$ from Eq. 6. The length is a measure of concentration of this distribution related to the circular variance (Batschelet, 1981). A long vector means more concentration (more coherent sketch orientations) and stronger differences from the average (long-term memory) distribution. Short vectors would result from sketch orientations that are similar to the long-term memory content.

8.1.2. Results

Orientation frequencies of the sketches of the ten downtown and four suburban interview locations are shown in Figure 29. For the near interview locations A, B, C, F and G, the south orientation was very prominent. For interview locations D, E, I and J, the orientations were distributed more equally. The distributions obtained at the near locations differed significantly from each other, $Chi^2(27, N = 254) = 88.036$; $p < 0.001$, indicating that recalled view orientation depended on interview location. For the distant locations, the south orientation was very striking and no differences between the histograms could be found.

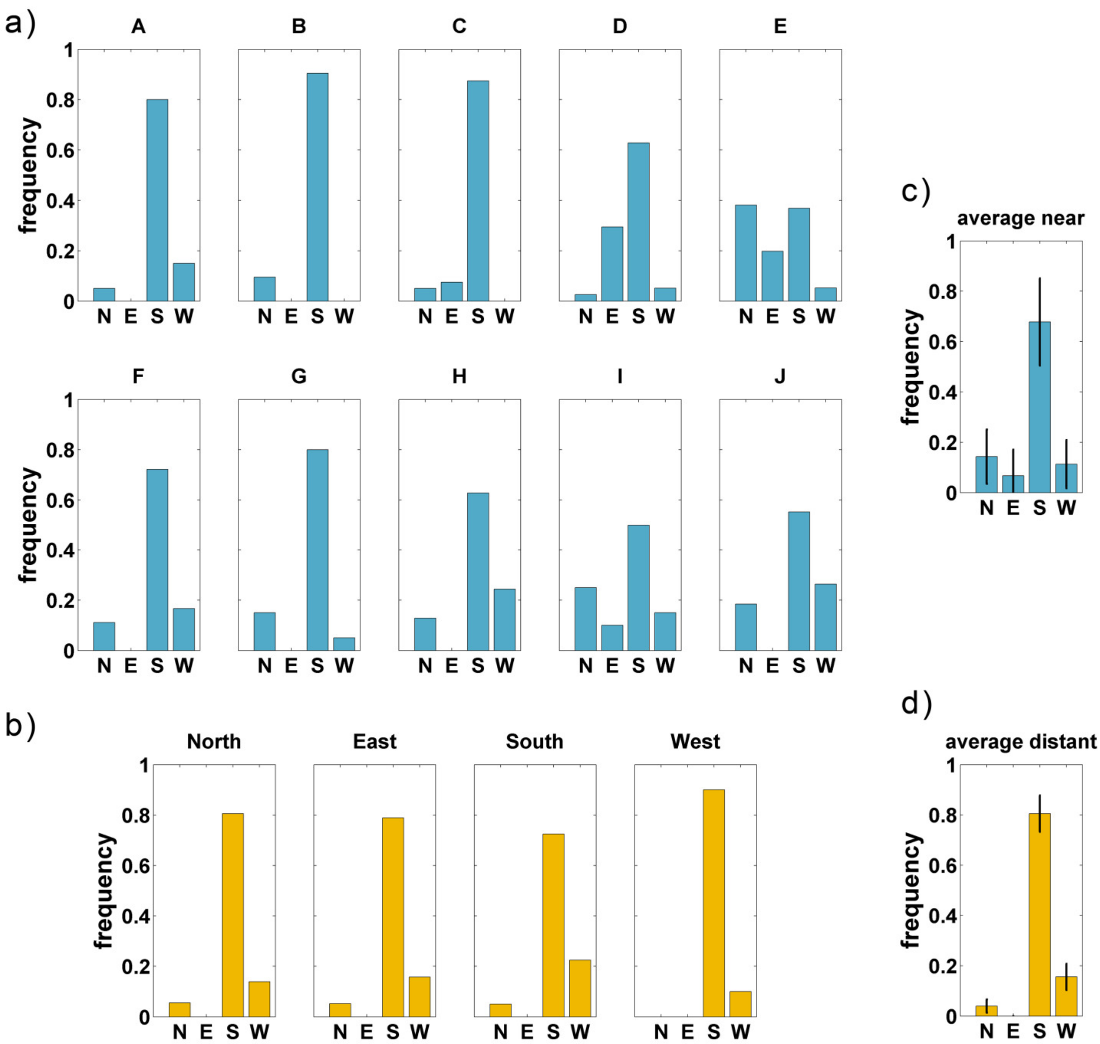

Figure 29. Sketch orientation frequencies for drawing the "Holzmarkt". a) Orientation frequencies of the near interview locations (A - J). The obtained frequencies differed significantly from each other. b) Orientation frequencies of the distant interview locations (north to west). c) & d) Average orientation frequencies with standard deviation of the near and distant condition. The y-axes show the frequencies of sketch map orientations, the x-axes the rated orientations (north, east, south, west).

The average distributions for near and distant interview locations are shown separately in Figure 29. These distributions were significantly different from each other ($Chi^2(3, N = 331) = 12.654$; $p < 0.01$) though comparable in shape. The orientation vectors obtained from the location-dependent components of the downtown interview locations (Eq. 7) are plotted in Figure 30, superimposed on a map of Tübingen, showing the target and interview locations. An overall tendency of the vectors to point to the target square was clearly apparent.

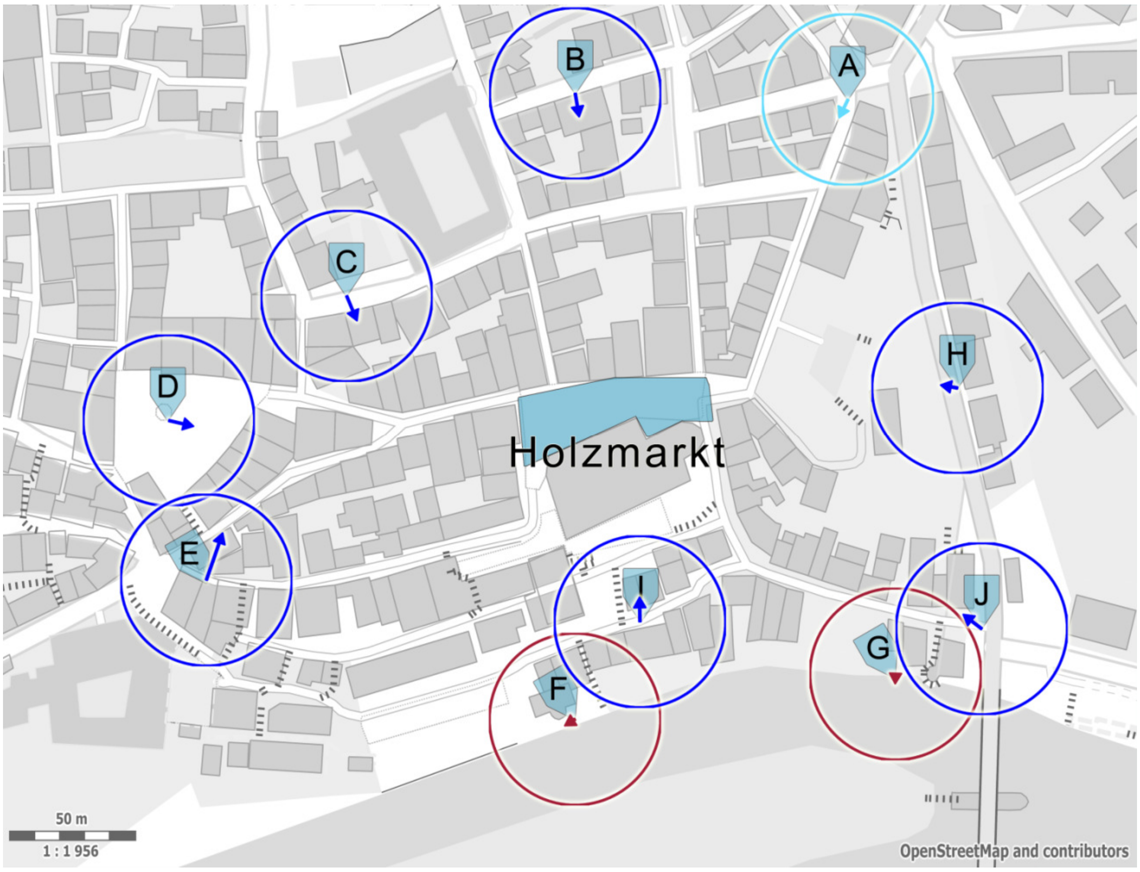

Figure 30. Downtown map of Tübingen with target square "Holzmarkt" and interview locations (A - J) for the near condition. The vectors show the average sketch map orientation at the respective interview location. At seven (blue circles) out of ten near locations, sketch orientations were found to point from the interview location in the direction of the target square (p < 0.05 or better). At one location, a strong tendency was indicated (A, cyan, p = 0.051). For two locations (F & G, red), no significant orientation effect could be found. Vector length reaches from zero to one (radius of circle) and is a measure of concentration of the location-dependent vectors. Map source: © OpenStreetMap contributors.

In order to test this tendency, the angular deviation between the location-dependent orientation vectors and the theoretical air-line vector obtained for each interview location by subtracting the coordinates of the target square (defined as the center of gravity of the blue area in Figure 30) from the coordinates of the interview locations was calculated (Eq. 3).

For each interview location, the deviation or bias of the data from a uniform distribution towards the theoretical direction was tested with a circular V-test (Batschelet, 1981; Berens, 2009), taking into account the vector length as a measure of concentration. The deviations towards the theoretical direction were significant ($p < 0.05$ or better) for seven out of ten interview locations (B, C, D, E, H and J; Figure 30), and marginally significant for an additional one ($p = 0.051$; location A in Figure 30). For two interview locations (F and G in Figure 30), no significant deviation from uniformity could be demonstrated.

Figure 31 shows the location-dependent vectors rotated such that the theoretical direction from each interview location to the target place appears in upwards direction. For this sample of 10 vectors, again, the circular V-test was applied, this time with the 0-degree-vector as a theoretical direction. For the overall sample, bias towards the theoretical direction was significant with $V(N = 254) = 0.234$; $u = 5.276$; $p < 0.001$.

For the four distant interview locations no such orientation effect could be found, $V(N = 77) = 0.038$; $u = 0.477$; $p = 0.317$.

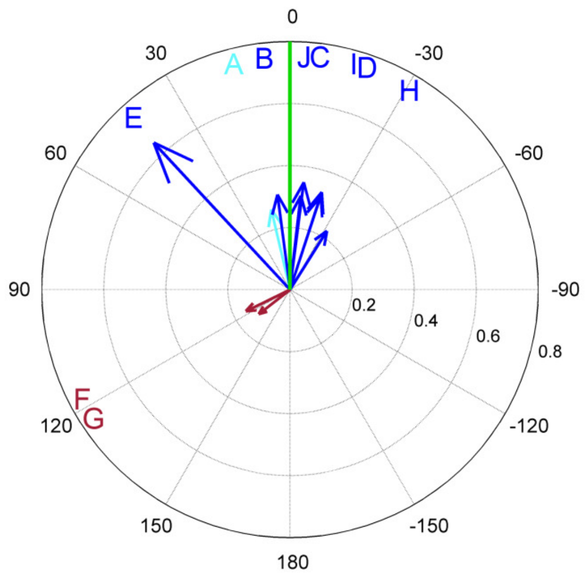

Figure 31. Location-dependent vectors from Figure 30, rotated to align the air-line directions from all near interview locations to 0 degrees (letters indicate interview locations A - J). Vectors are significantly biased towards the theoretical direction (green line, p < 0.001). Vector length reaches from zero to one and is a measure of concentration of the location-dependent vectors.

8.2. Experiment 4 b) - "Marktplatz"

To test the robustness of the findings of experiment 4 a) with respect to other target squares, another well-known square, the market square ("Marktplatz"), was chosen along with new interview locations and new participants and the previous experiment was repeated.

8.2.1. Material and methods

Eight new interview locations around the market square ("Marktplatz") were selected for the near condition [Figure 27 b) and Figure 56, green]. For the distant condition, the same locations as in experiment 4 a) were used except for the southern one, for which access could not be obtained again, resulting in three distant interview locations. 330 passers-by agreed to participate. The procedure was the same as in experiment 4 a). Examples of sketch maps are shown in appendix Figure 57.

Sketches were again categorized for orientation (north, east, south or west up) by three independent raters. From the 330 drawings, 306 were judged identically (93 %) with a chance-corrected inter-rater reliability of $\kappa = 0.93$. Only the 306 identically judged drawings were analyzed further (220 near condition, 86 distant condition). The mean age of the 306 participants (131 male, 175 female) whose maps were included was 37.4 years, their average years of residency in Tübingen was 12.7, their own judgment of local knowledge was 3.4 (with 1 = very poor and 9 = very good) and own judgment of how often they frequent the "Marktplatz" was 3.0, with 1 = very rarely and 9 = very often.

Average orientation frequencies for the near and distant conditions were calculated and subtracted from the histogram of the near and distant interview locations, respectively, yielding the location-dependent components of each distribution.

8.2.2. Results

Like in experiment 4 a), orientation frequencies of the sketches of the eight near interview locations [Figure 32 a)] differed significantly from each other ($Chi^2(21, N = 220) = 95.457$; $p < 0.001$). At the near interview locations C to H, there was a bias towards southern and western sketch orientations, while orientations where more evenly distributed at interview locations A and B. For the distant locations, no difference between the histograms could be found [Figure 32 b)]. There, mostly west and south sketch orientations were produced. Also, there was no significant difference between the near and distant average frequencies ($Chi^2(3, N = 306) = 3.986$; $p = 0.263$).

As shown in Figure 33, the majority of the location-dependent vectors of the near condition point towards the "Marktplatz" (center of gravity of green area in Figure 33). A significant bias of sketch orientations towards the air-line direction to the target square (center of gravity of the green area Figure 33) could be revealed by a circular V-test for six of the eight interview locations (A, B, E, F, G and H, Figure 33). The sample of eight location-dependent vectors, rotated to align their respective air-line directions, also showed a highly significant bias towards the theoretical direction at zero degrees, $V(N = 220) = 0.343$; $u = 7.203$; $p < 0.001$ (Figure 34). No bias could be detected for the three distant interview locations, $V(N = 86) = 0.099$; $u = 1.295$; $p = 0.098$.

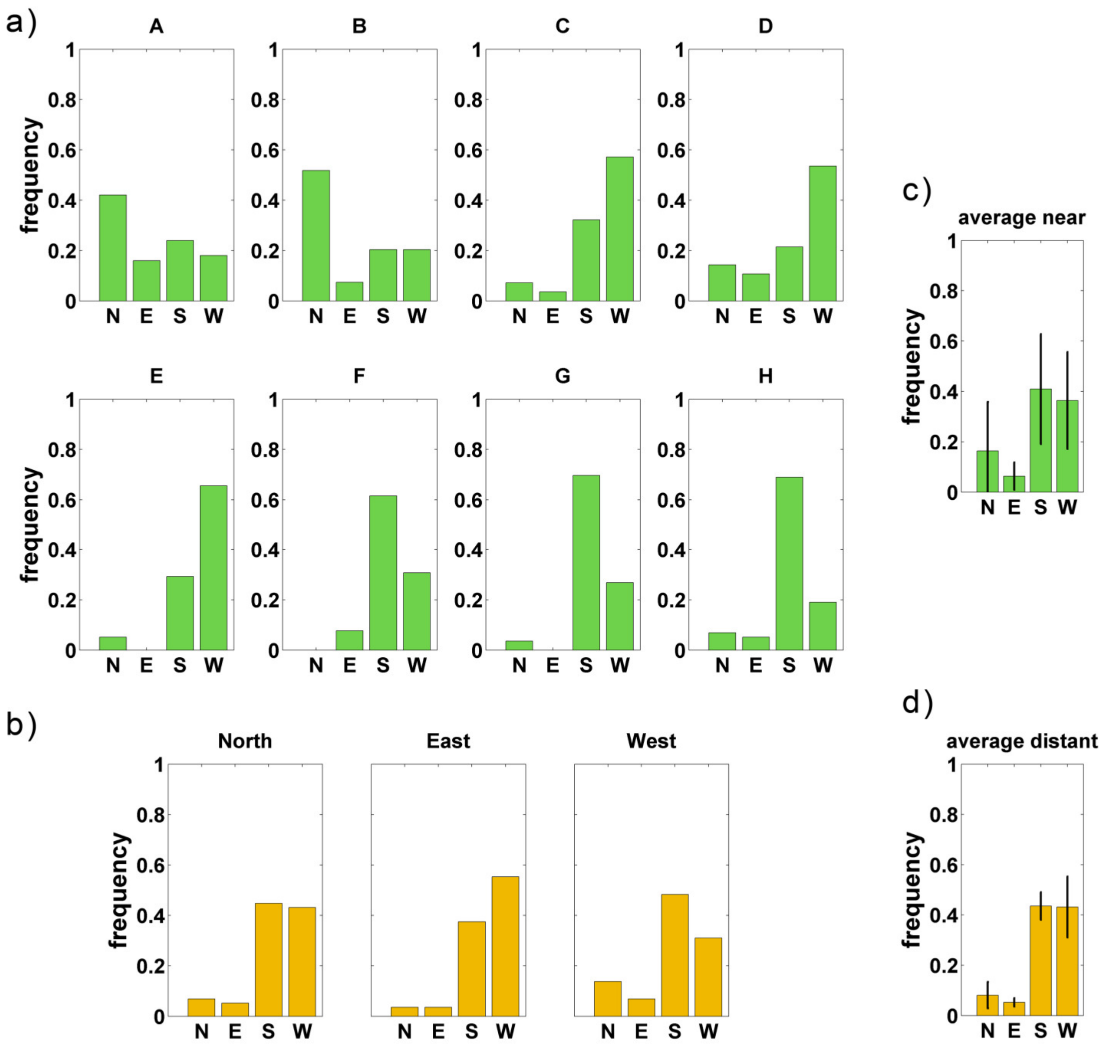

Figure 32. Sketch orientation frequencies for drawing the "Marktplatz". a) Orientation frequencies of the near interview locations (A - H). The obtained frequencies differed significantly from each other. b) Orientation frequencies of the distant interview locations (north, east and west). No significant difference could be found. c) & d) Average orientation frequencies with standard deviation of the near and distant condition, respectively. The y-axes show the frequencies of sketch map orientations, the x-axes the rated orientations (north, east, south, west).

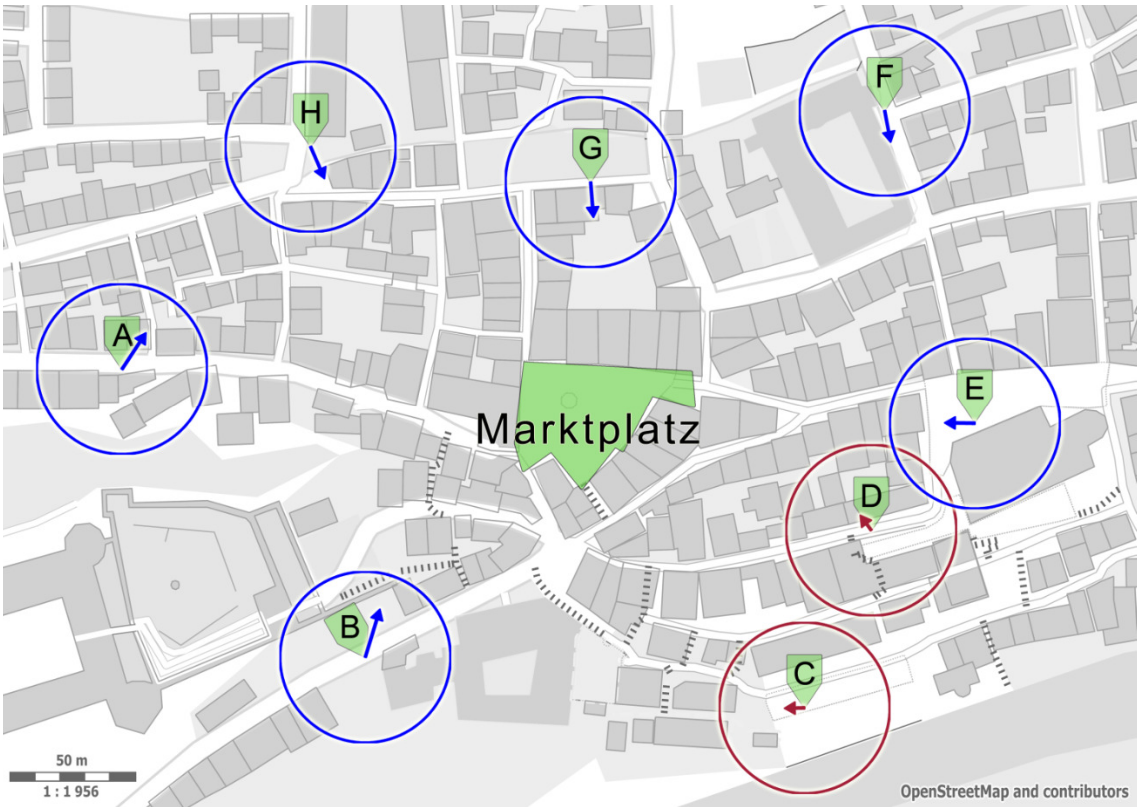

Figure 33. Downtown map of Tübingen with target square "Marktplatz", near interview locations (A - H) and location-dependent vectors drawn at these locations. Vectors at six (blue circles) out of eight interview locations point towards the target square ($p < 0.05$ or better). For two locations (C & D, red), no significant orientation effect could be found. Vector length reaches from zero to one (radius of circle) and is a measure of concentration of the location-dependent vectors. Map source: © OpenStreetMap contributors.

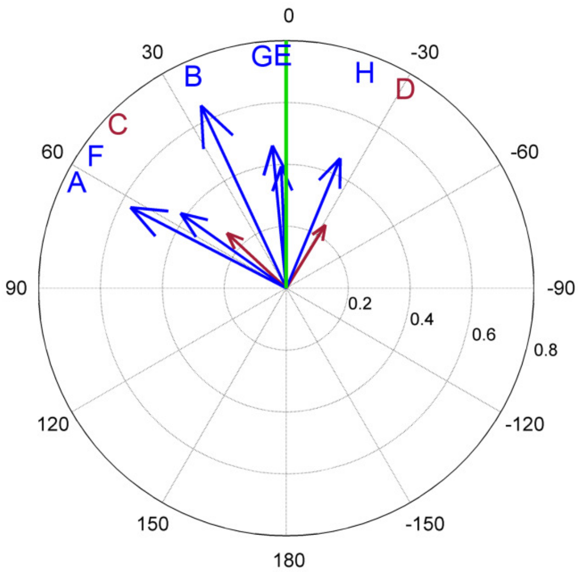

Figure 34. Location-dependent vectors from Figure 33, rotated to align the air-line directions from all interview locations to 0 degrees (letters indicate interview location A - H). Vectors were significantly biased towards the theoretical direction (green line, $p < 0.001$). Vector length reaches from zero to one and is a measure of concentration of the location-dependent vectors.

8.3. Discussion

The data presented in experiment 4 a) and b) indicate that visuo-spatial recall of out-of-sight places does not occur with a random or fixed orientation, but that recall orientation depends on both target and interview location.

The target square effect suggests a non-isotropic representation of each target square in long-term memory. For the distant (suburban) interview locations, orientation distributions were found that equal the average distributions taken over all near (downtown) locations. Therefore, it is concluded that the target square dependence is underlying all the measurements and is modulated by interview location-dependent effects visible only for the downtown interview locations. The average view distribution for the “Holzmarkt” square [experiment 4 a)] is strongly peaked with a “canonical view” in southward direction, depicting a landmark church building on top of a hillock. In contrast, the view distribution for the “Marktplatz” [experiment 4 b)] is more isotropic, probably reflecting the more balanced salience of the surrounding houses, although the town hall was still prominent and often drawn. These differences are probably related to the specific topography of each place. The “Holzmarkt” is rising to the south, with a prominent church building on top. Approaches from behind the church (Northwards) are almost impossible and very rarely walked. Drawings with the church on top might thus be favored by familiarity, alignment with environmental axes and the fact that uphill buildings will appear on top of the sketching paper. In contrast, the salience of the buildings surrounding the “Marktplatz” [experiment 4 b)] is much more balanced. The “Marktplatz” is also rising to the south, but the most prominent building, the city hall, appears not on top but on the Western side. Also, approaches from all directions are possible and frequently walked. Still, a peak in the experimental data towards “south” and “west” is apparent here, too. It is suggested that the long-term memory of either square is organized as a collection of discrete views [Figure 26 a)], sampling the various viewing directions with variable resolution much as has been suggested for view-dependence in face recognition (Bulthoff et al., 1995). Allocentric place memory might therefore be organized as a population code of orientation-specific memories. Indeed, neuronal specificities for views of places have been reported in the medial temporal lobe, see for example Ison et al. (2011) and Epstein (2008).

The formation of one or several canonical views of a place requires further study, concerning potential relationships to canonical views of landmark objects and the selection of one view or another as canonical. Reasons for selection might include: Distinctiveness to other places, availability and distribution of local landmarks, geometric layout, visual salience of objects, path options and functionality, or intrinsic axes of the environment (Mou & McNamara, 2002).

The distribution of recalled views depends also on interview location as was revealed by Chi-squared tests on the orientation histograms. For the near (downtown) interview locations, each local distribution is biased towards a preferred orientation, roughly corresponding to the air-line direction from the interview location to the target square. A view of the target square, oriented in the current approach direction, thus seems to be activated in a spatial working memory either by automated spatial updating when walking in the city or by a mental travel initiated when asked to draw the sketch or by both effects [see Figure 26 b) & c)]. Spatial updating itself could again be achieved by two mechanisms, either image transformation as discussed in view-based object recognition (Ullman & Basri, 1991) or by refreshing working memory from long-term memory.

In the introduction, a view-based model of spatial recall predicting that the directional distributions of recalled sketch maps are a mixture of a fixed long-term memory distribution and a set of position dependent working memory distributions (Eq. 5) was presented. As a direct test of this model, a maximum likelihood analysis was performed, assuming for the orientation histograms a multinomial distribution with four possible outcomes (N, E, S, W) and theoretical probabilities $\alpha l_k + (1-\alpha)w_{ik}$, where k numbers the four possible outcomes and $(l_1 \dots l_4)$ are the class averages over all interview locations, i.e. the assumed long-term memory contributions.

The log likelihood function reads

$$LL(\alpha) = \sum_{i=1}^{I}\left(c_i \sum_{k=1}^{4} n_{ik} \log(\alpha l_k + (1-\alpha)w_{ik})\right) \quad (8)$$

where n_{ik} is the number of orientations k found at interview location i, and the constant c_i is the logarithm of the multinomial coefficient for the local orientation distribution. Theoretical estimates for the working memory contributions at each interview location are derived from the local air-line directions φ_i (Eq. 3). The theoretical outcome probabilities for the assumed working memory distributions were set to $w_{i1} = c + 0.5\max(0, sin\,\varphi_i)$, $w_{i2} = c + 0.5\max(0, \cos\varphi_i)$, $w_{i3} = c + 0.5\max(0, -sin\,\varphi_i)$ and $w_{i4} = c + 0.5\max(0, -\cos\varphi_i)$, where $c = 1 - 0.5\,(|\sin\varphi_i| + |\cos\varphi_i|)$ is a constant, assuring that the four probabilities will add to 1. This distribution has the circular mean φ_i and variance 0.5, which reasonably approximates the location-dependent components shown in Figure 30 and Figure 33.

Figure 35 shows the relative log likelihood $LL(\alpha) - LL(\alpha^*)$ as a function of the mixing parameter α separately for the near and far interview locations in both experiments. For the "far" cases, the maximum likelihood estimator α^* is 1, i.e. adding position-dependent working memory contributions to the model does not improve likelihood in these cases. In contrast, for the "near" cases, the maximum likelihood estimates lie between 0.6 and 0.7; the horizontal lines in the plot are 99 % confidence intervals. A likelihood ratio test for $\alpha = 1$ vs. $\alpha < 1$ is significant with $p < 10^{-16}$ for the "near" cases in either experiment. The model with the location-dependent working memory component thus significantly improves the fit of the data.

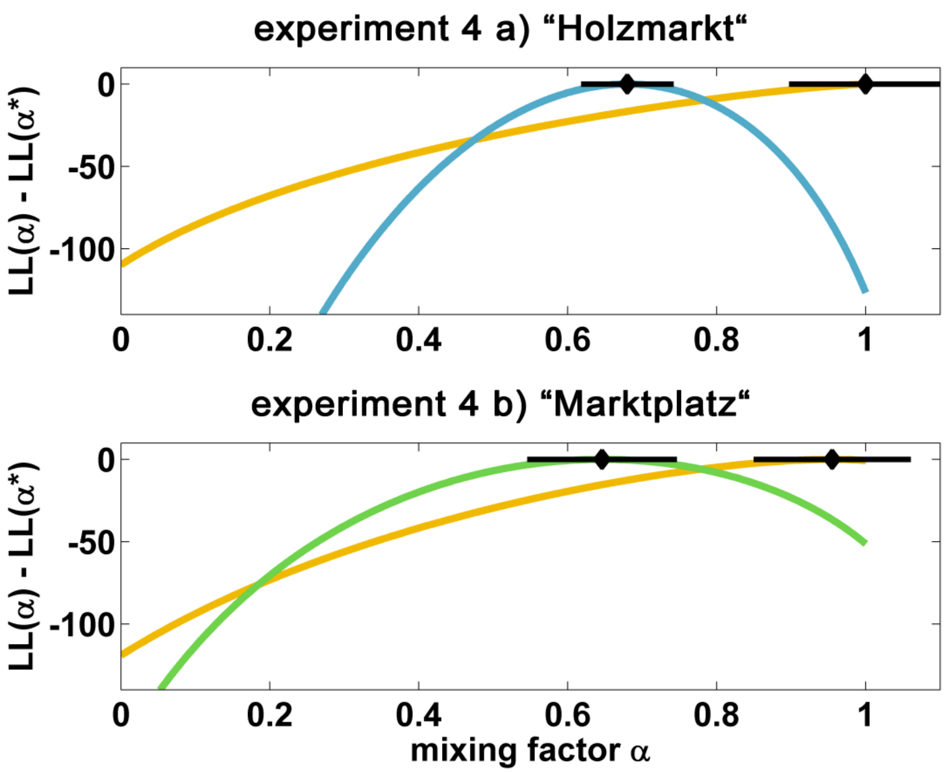

Figure 35. Likelihood analysis of the mixing model. Yellow: Distant locations; relative likelihood peaks for α = 1, indicating that orientation distributions do not depend on air-line direction. Blue ("Holzmarkt") and green ("Marktplatz"): Near locations; relative likelihood peaks at α* < 1, indicating the orientation distributions do depend on air-line direction in this condition. The black markers indicate α* with 99 % confidence intervals. The y-axes show the relative log likelihood LL(α) - LL(α*), the x-axes the mixing factors α for working and long-term memory contributions.*

It cannot be decided from the data whether recall bias is strictly towards the air-line direction or towards the actual entry view obtained when walking to the target place along the street network; although, in a view-based account, the latter seems more plausible. This will be investigated in experiment 5 (see below). Indeed, this might have been the problem with the interview location D in experiment 4 b) from which two roughly equidistant routes to the target place exist, each with opposite entry directions into the target square.

No location-dependent effect was found for the distant (suburban) interview locations; concluding that in these cases recall did not depend on working memory processes such as spatial updating or mental travel. Of course, other working memory effects might still be involved. Since only two distance conditions were used, downtown and suburban, one cannot decide how far the location-dependent effect extends around the target place or if there is a gradual decay as could be modeled by a distance-dependent factor α in Eq. 4. This will be investigated in experiment 5. It is clear, however, that the effect extends over tens to hundreds of meters which seem to be included in spatial working memory.

Another parameter in addition to the mere distance could be regionalization and spatial hierarchies. In virtual environments, navigators were shown to prefer routes that cross fewer region boundaries over equidistant routes through multiple regions (Wiener & Mallot, 2003). In this experiment, regions were defined by the semantic class of landmark objects. In a pointing experiment, Wang & Brockmole (2003) demonstrated that information from nested environments may be kept separate in spatial representations. In the city environments used in the present study, there are various configurations of buildings, roads, shops, etc., which segregate the environment into quarters, districts, neighborhoods, etc. Therefore, it seems possible that the extension of spatial working memory is defined by region boundaries rather than by metric distance. This might also explain the results for the interview locations F and G in experiment 4 a) and C in experiment 4 b): They were probably attributed to the region “riverfront” and not “downtown”, and therefore no or only weak connections to the target places existed while the experiment took place.

The presented theoretical account is clearly able to explain the data. In addition, the findings by Basten et al. (2012) on view-based priming of recall by mental travel also fit into the overall scheme. In that study, all interviews were carried out at a distant location [between the north and west location in Figure 27 a)], and simple recall of the “Holzmarkt” square revealed the same view preference reported here for distant interview locations. Mental travel across the “Holzmarkt”, however, primed view-specific recall in the direction of travel, indicating that mental travel, just as actual walking in downtown Tübingen, activates view-specific working memories.

Alternative models of spatial working memory not based on views but on object representations and maps have been presented by Byrne et al. (2007), Tatler & Land (2011) and Loomis et al. (2013). While the data here do not strictly rule out these models, they make clear that representations of places are not unique entities that are always activated in their entirety, but that parts of place representations can play independent roles in spatial recall. Such parts are oriented and have therefore been referred to as “views” in these experiments. Alternatively, such parts could be landmarks or houses located at one side of a square, or names or other properties of such landmarks or houses, as might have been the case in the experiments reported by Bisiach & Luzzatti (1978). The considered parts of place representations are view-like in two respects: First, the target square effect (canonical view) shows that oriented parts of a place representation can be anisotropically distributed. Second, priming by spatial nearness (physical and imaginary) activates oriented parts of the representation of places, not place representations in their entirety. This finding is in line with previous results of Mallot & Gillner (2000) who showed that associative landmark usage depends on oriented parts of place representations rather than on representations of entire places. Overall, oriented “views” are suggested to form a separate level of granularity in spatial representation that can be activated whenever view-specific information is required.

To further investigate the nature of these “views”, experiment 5 was carried out in which participants were presented with pictures of different locations that were taken at eye level and therefore possibly qualify as views that could be stored in long-term memory.

8. Experiment 4 - View-based spatial memory

9. Experiment 5 - View recognition of places

Navigating in natural or virtual environments requires assessment of the visual scene and acquisition of new objects and landmarks as well as comparisons with memorized features of familiar places. Therefore, interactions between working memory and long-term memory must take place. In experiment 4, view-based organization of these memories was made a subject of discussion. Thereby, spatial long-term memories are thought to be stored as “views” of memorized places. Different views are connected with each other with respect to their spatial location and spatial connection(s). That is, if two neighboring places are connected through a certain path, and this route is known, also the views representing these places and the view that leads from one to the other place are connected.

In experiment 4, participants were asked to draw sketches of a certain place at different interview locations, and it turned out that these drawings were aligned towards that place - without instructing the participants to do so. However, this only happened at nearby interview locations and not at distant ones. In this fifth experiment, the view-based concept is enlarged upon. Therefore, it was investigated whether this alignment towards the places could also be found by presenting pictures of various views of different downtown places while being in the area, similar to experiments 4 a) and b). These views are assumed to represent - or at least activate - the stored (long-term) memory content, which then gets transferred into working memory. It is hypothesized that pictures that show the view of a place that would be seen upon walking directly from the interview location to this place would be recognized faster than other views because of a pre-activation or priming of these or similar views in memory.

9.1. Material and methods

9.1.1. Participants

Passers-by at seven locations (see Figure 37) in Tübingen (Germany) were approached during day time and asked “if they would participate in a quick interview for a navigational study” which would take approximately three to five minutes. 140 passers-by (72 male, 68 female; 20 at each interview location) participated in the study with an average age of 28.2 years (see Table 1). They were informed about the type of the collected data and the general procedure. Only residents of the city of Tübingen with a residency of one year and more could participate. This was done so that a minimum of local knowledge could be assumed. Participants were not asked for their names and accordingly were not required to give their consent in written form. Participants were free to terminate their participation at any time, simply by walking away. The informed consent procedures adhere to the guidelines of the Declaration of Helsinki, approval by the local ethics committee was not required.

9.1.2. Setup

Participants were shown previously taken color photographs (an example is shown in Figure 36) of first person views of the city of Tübingen on a tablet (Venue 11 Pro, Dell). They were asked to identify them as the market place or not ("Ist das der Marktplatz?") by pressing one of two buttons (yes / no) that were also displayed on one side of the tablet's display, depending on the handedness of the participants (see Figure 36). For right-handed people, the buttons were on the right side and for left-handed people on the left side of the screen. The vertical position of the yes and no buttons was varied randomly between participants.

The display size of the tablet measured 10.8 inches in diagonal with a resolution of 1920 x 1080 pixels. The displayed pictures had a resolution of 1630 x 1080 pixels. The buttons had a size of 290 x 540 pixels each and covered about 20 % of the screen.

Figure 36. Schematic view of the tablet and displayed stimuli. On the tablet, the photographs were displayed on one side and the interaction buttons on the other side. The sides were determined by the handedness of the participants before the experiment. The picture here shows a photograph of the "Holzmarkt", an adjacent place of the "Marktplatz". Therefore, participants were expected to press the 'no'-button ("Nein") here as it is not the "Marktplatz".

9.1.3. Locations and stimuli

Participants were shown photographs of nine different places and squares in the city center of Tübingen at seven interview locations around town (see Figure 37). For each square, at least two photographs from different viewpoints and viewing angles were taken, depending on the size of the square and number of avenues leading to and away from the place. The photographs showed typical views and features of each place so that the participants should be able to easily identify and differentiate between them. For the target square - the market square ("Marktplatz") - nine different photographs were taken, at least one at each avenue, with views towards the square (see appendix Figure 58 and Figure 59).

There were six interview locations (B, C, D, E, F and M) downtown and one (A) about 2 km away as control condition at which passers-by were asked for participation. These six downtown interview locations were at the same places of which photographs have been taken and which were shown during the experiment.

Figure 37. Interview locations and target square. The blue and green markers and letters mark the six downtown locations at which passers-by were interviewed. The market square is colored in green; its four avenues (NW, NE, SE, SW) are marked with red circles. Marker A for the distant interview location is out of the downtown area and not shown here. Map source: © OpenStreetMap contributors.

9.1.4. Procedure

Passers-by were approached at the interview locations and asked for participation (see above). Care was taken to approach participants walking in different directions. Approach was from sideways with respect to the participant's heading. Upon being approached, they stopped but did not change their general body orientation. They were further asked for their residency and handedness; subsequently, the experimental program was started on the tablet by the interviewer who also entered the data at this point (handedness and interview location). The tablet was handed over to the participants and the experiment started upon button press.

First, participants were presented with three training trials that were excluded from analysis so that they could familiarize with the tablet and the procedure. In these training trials, different pictures of places around Tübingen were shown which were not used again in the main test trials of the experiment. The participants were informed about these training and test phases by visual display.

Then, the main test began in which participants had to solve 90 trials, which took them approximately 5 minutes. Later, the number of trials was reduced to 58 trials, which then took approximately 3 minutes. This was done after a few participants have been tested because they seemed to lose attention towards the end of the experiment and they asked “how long the experiment still would take” and if it would end soon. Afterwards, the participants were shown the test result of correct answers in percent. This was done to reward them for their participation and also to encourage them if they came in groups.

The 90 and 58 trials started and ended with several flanker trials, which were also excluded from analysis. Consequently, 81 and 45 trials were analyzed. Out of the 81 trials, 27 showed pictures of the target square (market square) with each picture repeated three times. Each non-target picture (views of other places) was repeated two times. Later, when only 45 trials were used, pictures of the market place were shown twice and non-target pictures once. The picture of each trial, that means the sequence of pictures, was pseudo-randomly selected. Thereby, the same picture could not occur twice after another as well as different pictures of the same place could not occur more than two times in succession.

After the experiment, participants were asked for their age, years of residency in Tübingen, local knowledge on a scale from one (very poor) to nine (very good), estimated walking duration from the interview location to the market place, whether they visited the market place during the last 30 minutes and whether they planned to visit the market place within the near future. For the distant interview location, only the age and years of residency were acquired (see Table 1).

Table 1. Information on the participants. Questions concerning the market place asked at the market place (M) were excluded from the mean and standard deviation (data in brackets). Interview location A was the distant control condition.

interview location	**mean age (years)**	**local knowledge (1 - 9)**	**years of residency**	**estimated walking duration (min)**	**already visited market place (%)**	**planned to visit market place (%)**
A	22.2	---	3.25	---	---	---
B	34.0	7.25	11.40	3.55	25	15
C	31.1	6.60	13.30	3.70	45	25
D	26.9	7.25	11.30	1.83	20	5
E	25.7	6.75	8.15	1.90	15	5
F	29.5	6.35	7.14	2.85	35	15
M	28.1	5.50	10.75	(0)	(100)	(100)
mean SD	**28.2 ± 3.8**	**6.62 ± 0.65**	**9.33 ± 3.39**	**2.77 ± 0.88**	**28.0 ± 12.0**	**13.0 ± 8.4**

9.2. Results

Participants' performance on identifying the target square was very good with a hit rate of 98.6 % and a correct rejection rate of 94.7 %, leading to a sensitivity index (d') of 3.827 (Figure 38). These levels were comparable at all interview locations (see appendix Table 4).

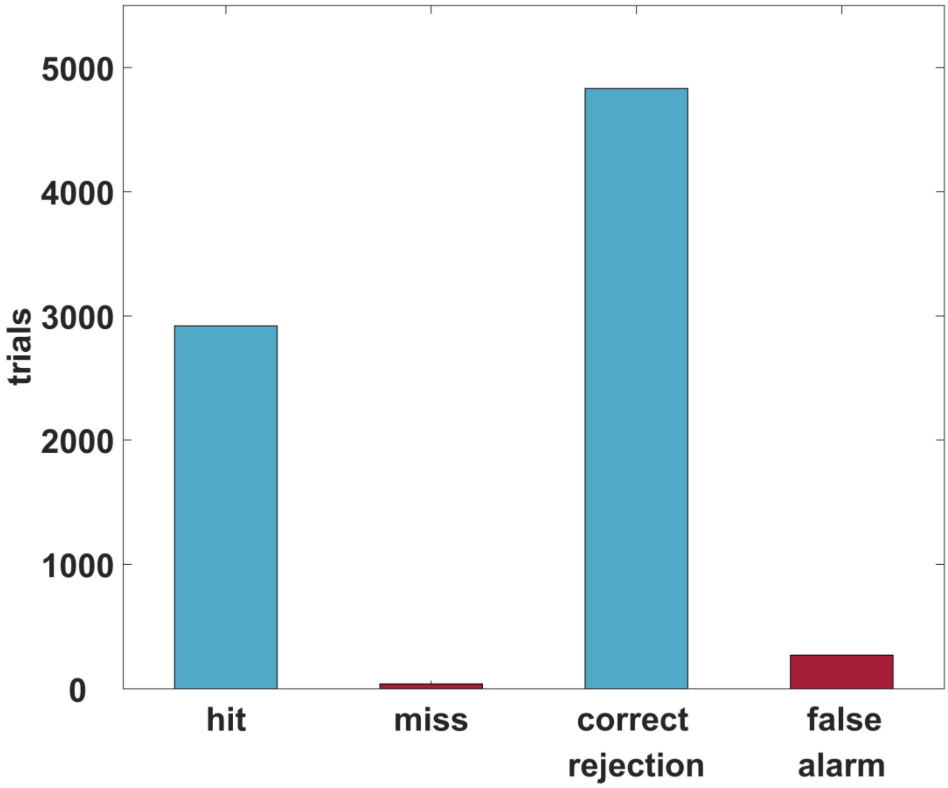

Figure 38. Participants' signal detection rate on photographs. Participants were very good in discriminating between target and non-target photographs. The y-axis shows the number of trials.

The average response time of all participants at all interview locations was 1963.8 ± 866.7 ms. A closer look at the response time distribution revealed several extremes: Four values were extremely small (down to 62 ms). It is unlikely that the participants were able to look at the picture, make a judgement about the location and push the button in such a brief time. These trials were excluded from further analysis and a lower limit of at least 500 ms in response time was introduced. There were also trials in which extremely large response times up to 45 s were detected. Again, these values are not likely to represent the required time to solve the task. There were also reports of cases in which a button was pushed on the touchscreen of the tablet but it was not detected (and had to be pushed again leading to larger response times). Therefore, an upper limit of 5000 ms response time (about 3.5 times the standard deviation) was introduced, too. With both limits, 328 trials in total had to be excluded from further analysis (about 4 %). There were 7736 trials remaining with an average response time of 1692.3 ± 449.5 ms.

Response times were analyzed for differences between genders (Figure 39). There was no significant difference in response time between male (1715.6 ± 431.6 ms) and female (1667.6 ± 469.6 ms) participants, $t(138) = 0.630$; $p > 0.05$.

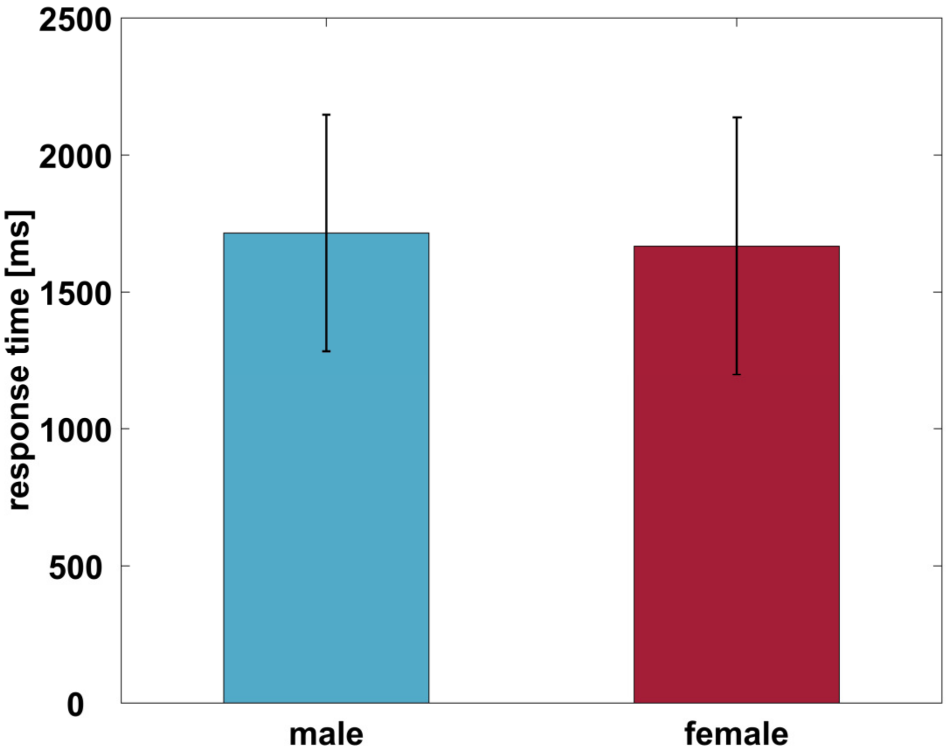

Figure 39. Effect of gender on response time. No significant difference in response time was present. The y-axis shows the response time in ms, the x-axis the participants' gender. Error bars show the standard deviation.

Response times were subsequently analyzed according to the interview location and the presented target place pictures; that is, only trials in which the market place was identified correctly were considered (Figure 38, first bar "hit").

A two-way repeated measures ANOVA of interview location and market place picture indicated a significant main effect of interview location, $F(6, 125) = 2.776$; $p < 0.05$ (Figure 40), and also of the presented market place picture, $F(6.80, 850.21) = 5.929$; $p < 0.001$ (Figure 41). There was no interaction between these two factors. Since Mauchly's test indicated that the assumption of sphericity had been violated for market place picture, Greenhouse-Geisser corrected values were used. A Bonferroni corrected post-hoc analysis revealed a significant difference between interview location A and M ($p < 0.05$, Figure 40).
Response times were smallest at the distant interview location with 1343.6 ± 316.6 ms on average (A, yellow). Contrary, participants interviewed at the target location "Marktplatz" (M, green) had the largest response times with 1757.8 ± 409.0 ms on average. For the pictures of the market place, a variation between the pictures was present but without any striking deviations (Figure 41).

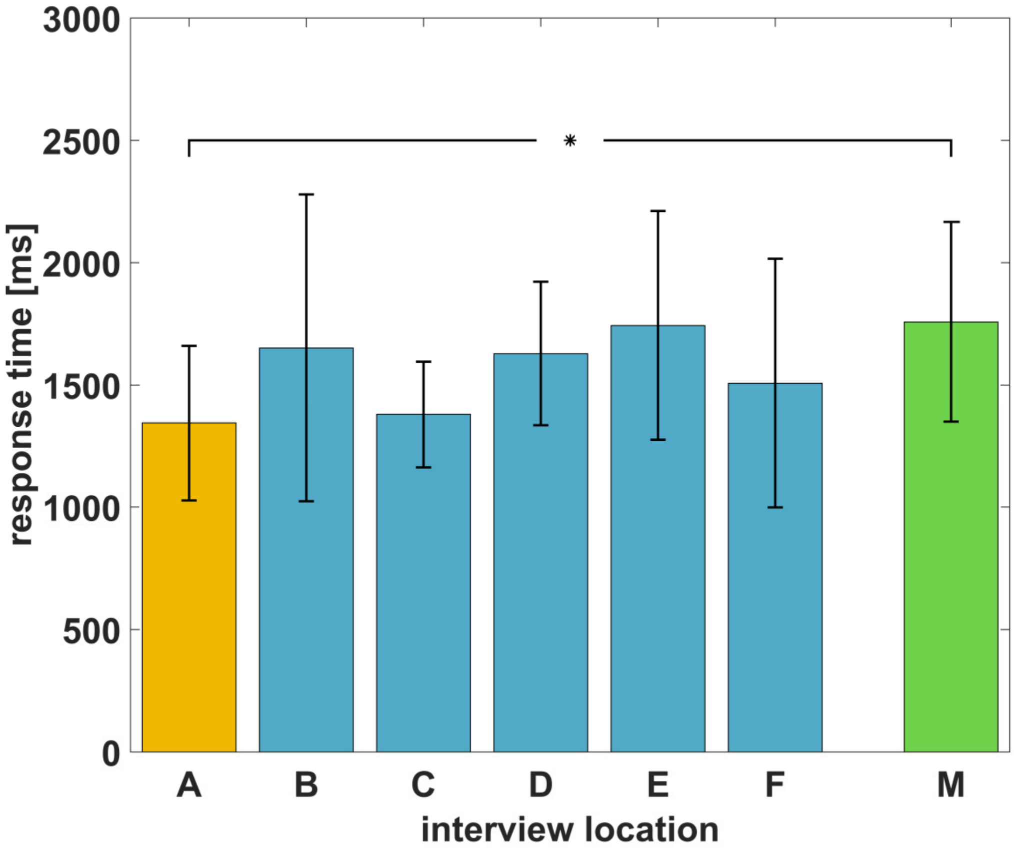

Figure 40. Average response time per interview location for target place pictures. Response times were significantly different for the interview locations. The y-axis shows the response time, the x-axis the interview locations. Error bars show the standard deviation.

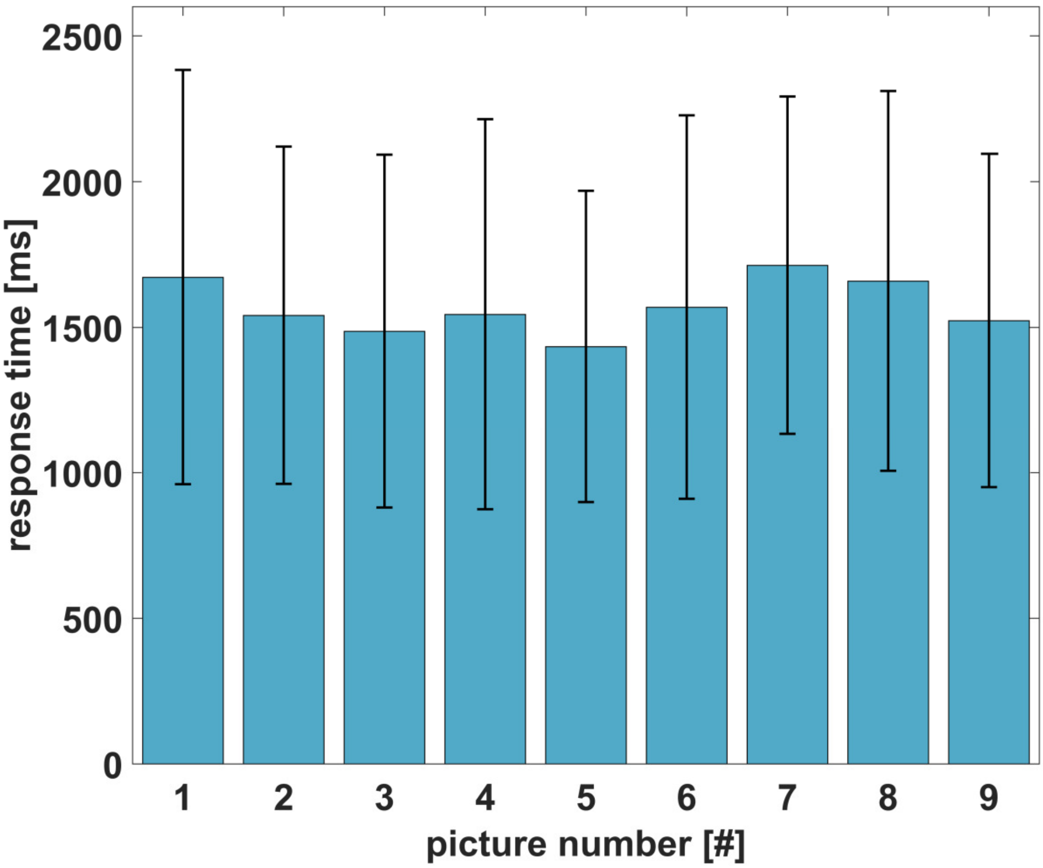

Figure 41. Average response time per target place image. Response times varied depending on which image of the market place was shown. The y-axis shows the response time in milliseconds, the x-axis the number of the market place picture. Error bars show the standard deviation.

An analysis of response time for interview location and picture was also done for all photographs, which includes pictures of the market place and additional pictures of other squares in town (see material and methods). Again, a two-way repeated measures ANOVA of interview location and picture was conducted, indicating a significant main effect of interview location, $F(6, 130) = 2.866$; $p < 0.05$ (see appendix Figure 60), and also of the presented picture, $F(6.60, 858.27) = 20.978$; $p < 0.001$. There was no interaction between these two factors. Since Mauchly's test indicated that the assumption of sphericity had been violated for picture, Greenhouse-Geisser corrected values were used. The response times of this analysis showed a comparable pattern to those of the market place pictures only, however, response times were increased in general.

In the next step, the photographs' headings of market place images were grouped and their means analyzed according to the cardinal directions north, east, south and west. The nine images taken at the market place were distributed as follows: One picture in group "north", four pictures in group "east", three pictures in group "south" and one picture in group "west" (see also appendix Figure 58).

A two-way repeated measures ANOVA of interview location and cardinal direction was carried out, indicating highly significant main effects of interview location, $F(6, 132) = 5.341$; $p < 0.001$, and cardinal direction, $F(2.60, 343.69) = 17.818$; $p < 0.001$. There was also an interaction between cardinal directions and interview location, $F(15.62, 343.69) = 1.955$; $p < 0.01$, indicating that the recognition of views of a place is affected by the interview location as it was hypothesized.

Greenhouse-Geisser corrected values were used here as a violation of the assumption of sphericity was indicated for cardinal direction and interaction by Mauchly's test. The results are depicted in Figure 42. Bonferroni corrected post-hoc tests showed significant differences ($p < 0.05$) between north and west as well as east and west at interview location A. At interview location C there was a significant difference ($p < 0.05$) between east and west. At interview location M, a highly significant difference ($p < 0.001$) was found between south and west as well as a significant difference ($p < 0.01$) between east and west. For interview locations B, D, E and F, no differences could be found.

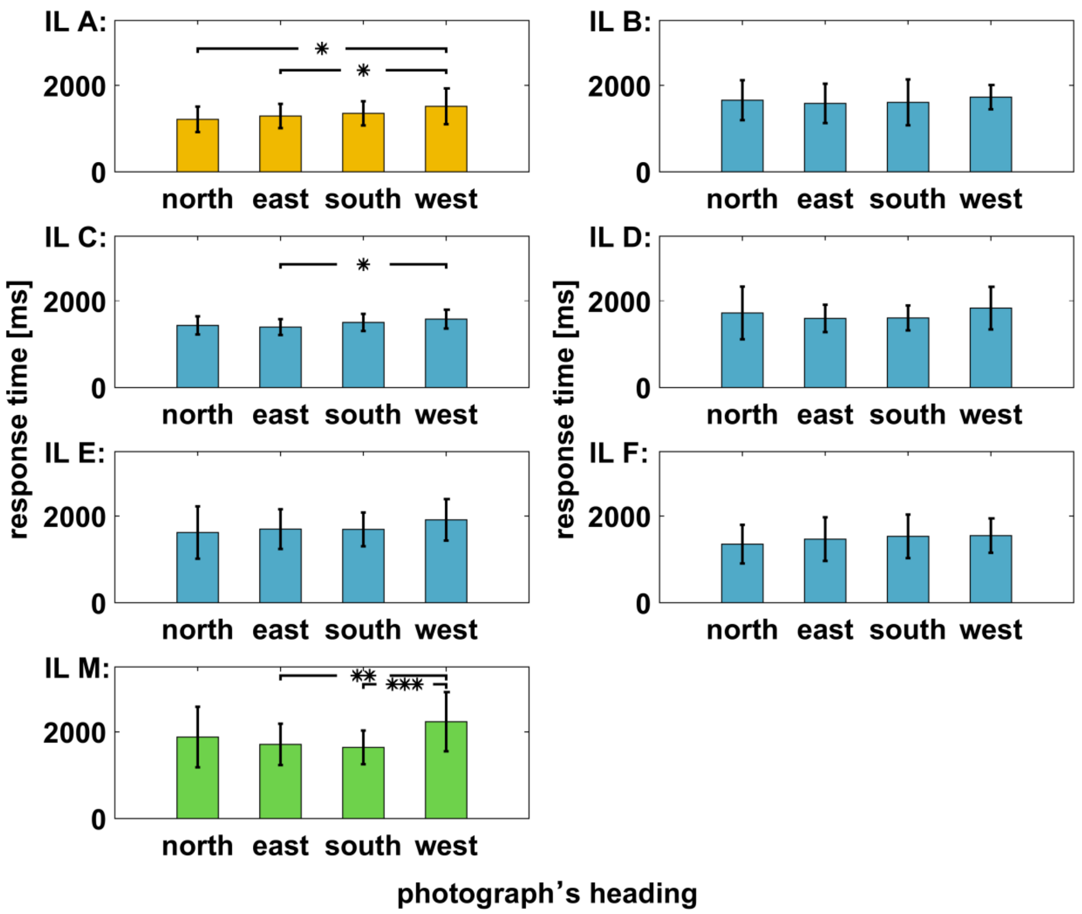

Figure 42. Response times per cardinal direction for market place photographs. Response times differed depending on interview location (IL) and photograph headings. The y-axes show the response time in milliseconds, the x-axes the photographs' cardinal headings. Error bars show the standard deviation.

Again, this analysis was done for all photographs (pictures of market place and other squares in town). The 36 images taken at all places were distributed as follows: Eleven pictures in group "north", ten pictures in group "east", seven pictures in group "south" and eight pictures in group "west". A two-way repeated measures ANOVA of interview location and cardinal direction indicated a significant main effect of interview location, $F(6, 133) = 3.244$; $p < 0.01$, and of cardinal direction, $F(2.71, 360.33) = 54.127$; $p < 0.001$. There was also an interaction between both factors, $F(16.26, 360.33) = 2.627$; $p < 0.01$, which indicates that the hypothesized effect is also present for (nearby) non-target places. Greenhouse-Geisser corrected values were used here as a violation of the assumption of sphericity was indicated for cardinal direction and interaction by Mauchly's test.

As shown in Figure 43, significant differences in response times between the four cardinal directions at each interview location were indicated by Bonferroni corrected post-hoc tests. At interview location B, highly significant differences ($p < 0.001$) were found between north and east, north and south as well as north and west. At interview location C, highly significant differences ($p < 0.001$) were found between north and east, east and south as well as east and west. A difference with $p < 0.01$ was present between north and south. Between north and west, a significant difference ($p < 0.05$) could be found, too. At interview location E, significant differences ($p < 0.01$) were found between north and east as well as north and south. At interview location F, differences between east and west were significant with $p < 0.05$. Finally, at interview location M there were highly significant differences ($p < 0.001$) between north and east as well as north and south. A significant difference with $p < 0.05$ could be found between north and west.

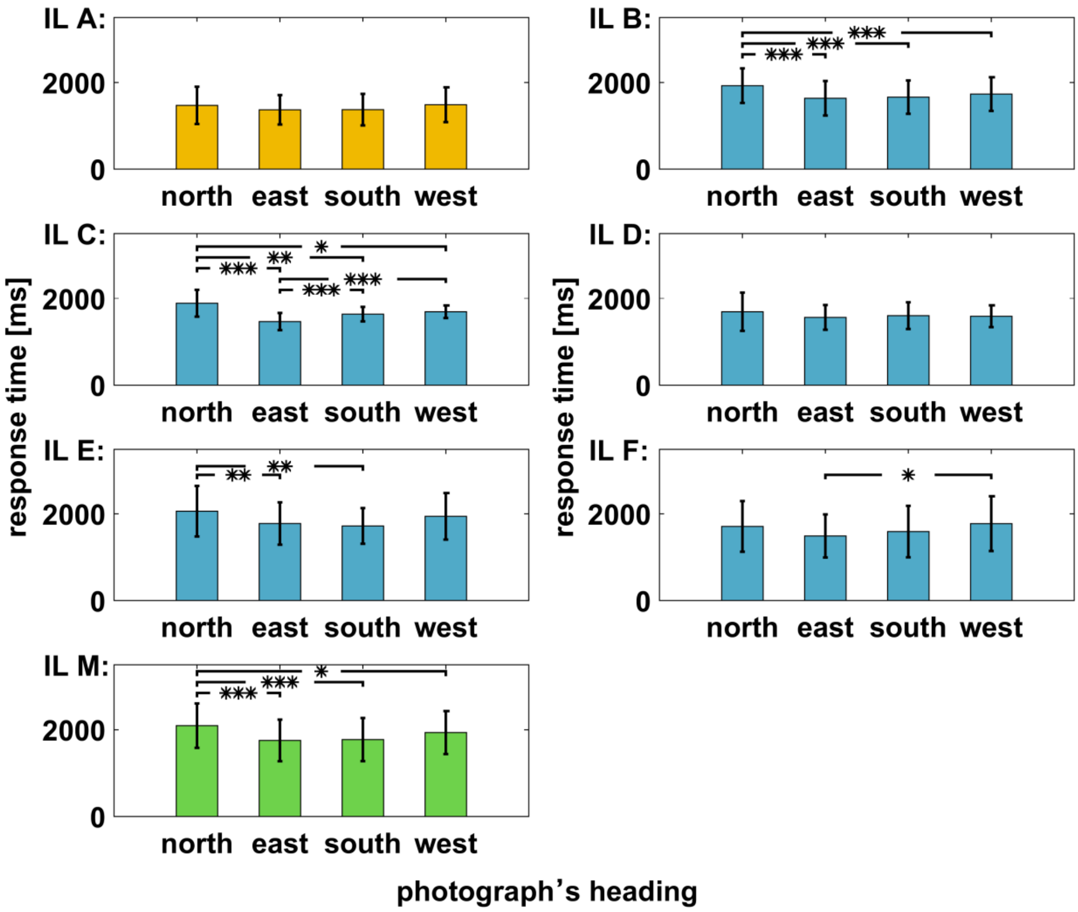

Figure 43. Response times per cardinal direction for all photographs. Response times differed depending on interview location (IL) and photograph headings. The y-axes show the response times in milliseconds, the x-axes the photographs' cardinal headings. Error bars show the standard deviation.

To further investigate the interaction between interview locations and photographs, data plots with response times for every interview location and picture were made. Figure 44 a) shows the average response times (blue lines) for the nine pictures of the target place of participants interviewed at locations D and F [as an example; for the other interview locations see appendix Figure 61 a)], which were drawn on an outline of the target place “Marktplatz” (dark lines). The position in the corners and angles of the lines reflect the position and direction the pictures were taken in, the length of the blue lines indicates the response time (short lines - small response times, long lines - large response times). One can see that most of the lines and therefore response times were longer in Figure 44 a) at interview location D compared to interview location F. However, it is not clear if there is any systematic difference among the different pictures depending on the interview location.

Therefore, the average response times of all participants, nearby interview locations (= locations B to F) and pictures were taken and subtracted from the individual response times, analogous to experiment 4, and multiplied with minus one. This method not only highlights the differences but also eliminates any general directional bias as it was found in experiment 4. The results are shown in Figure 44 b).

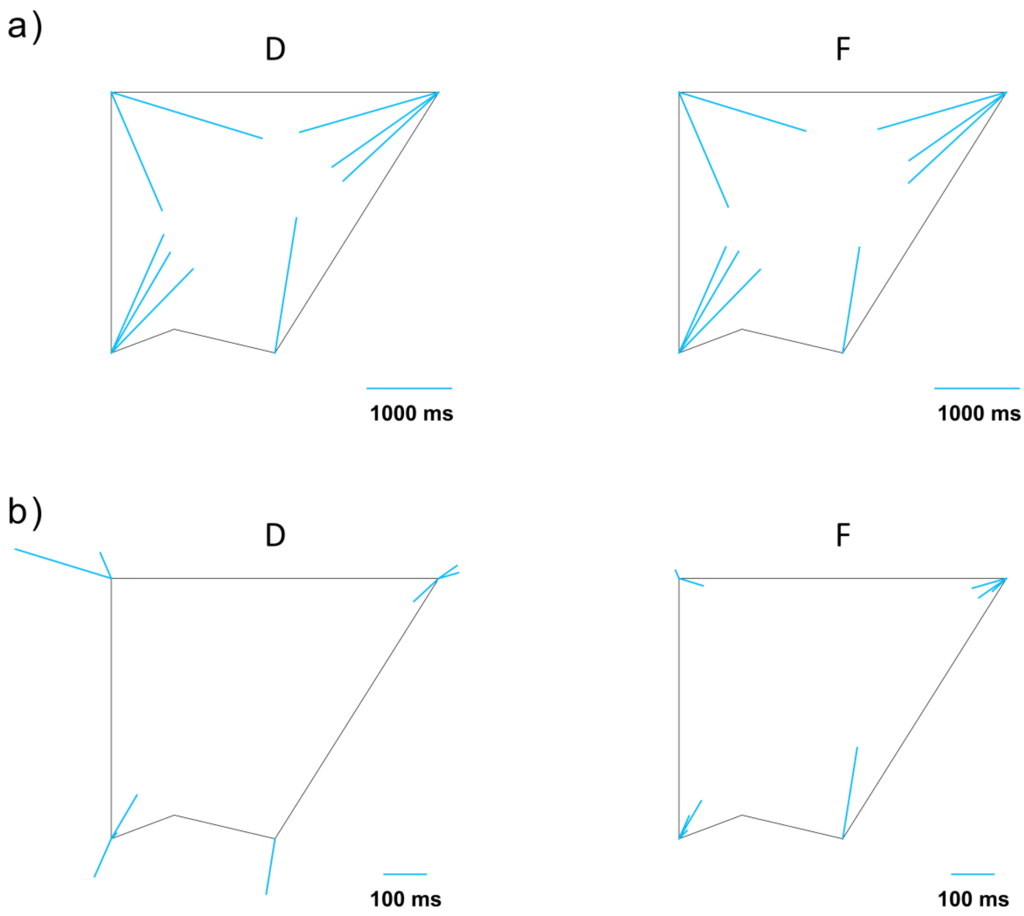

Figure 44. Response times and response time differences at two interview locations. a) Typical response times at interview locations D and F. Participants were faster at interview location F as shorter lines denote. Dark lines indicate the outline of the target square (north up). Blue lines indicate view point and angle of presented photographs of the target place and their length indicates response times. b) Response time differences after subtraction of the mean. Participants at interview location F were typically faster than average, while participants at interview location D where mostly slower than average. Here, blue lines indicate view point and angle of presented photographs of the target place and their length response time differences. Lines pointing inwards indicate shorter response times, lines pointing outwards show responses that took longer than average [Please note that the scale in a) is 1000 ms, while in b) it is 100 ms].

Lines pointing to the inside of the sketched market place indicate responses that were faster than the average response time. Lines which point to the outside show responses that were slower than average. In Figure 44 b) at interview location F, the blue line on the lower right side points inside and is longer than most of the other lines, indicating very small response times for the picture taken at this position compared to the other pictures. For pictures taken at the other three entryways, response times rather point inward. At interview location D, a mix of both inward and outward lines was found [see Figure 44 b)]. The response times of interview locations A and C all point inward, though with different lengths [see appendix Figure 61 a)]. In turn, response times that were collected at interview locations B and E rather point outwards.

Still, it is difficult to see if there are systematic deviations at a given interview location. From experiment 4 and the analysis of this experiment (see Figure 40) it is known that at distant interview locations a different route planning mechanism takes place and responses were faster. In contrast, the closer the interview location was to the target place, the longer it took the participants to respond. The subtraction of the average response time of all interview locations eliminated a general directional bias (as has been found in experiment 4) but also covers up subtle response time differences, especially at those interview locations with fast responses (and therefore small numbers). A two-way repeated measures ANOVA of interview location and individual market place picture, which was the first ANOVA in this experiment 5 (see above), indicated significant main effects but no interactions. On the other hand, both ANOVAs of interview location and cardinal directions of market place images and cardinal directions of all images revealed interactions between these factors. To analyze whether there is an interaction with individual images, the two main effect factors (averages of interview location and picture) need to be subtracted out from the response times:
Let the average response time of a certain market place image i at a certain interview location j be t_{ij}. Let further the average over all interview locations per market place image be denoted as $\bar{t}_{i\cdot}$ and the average over all market place images per interview location be $\bar{t}_{\cdot j}$, where $i = 1$ to 9 and $j =$ B to F. The average response time over all interview locations is calculated as

$$\bar{t}_{i\cdot} = \frac{1}{5}\sum_{j=B}^{F} t_{ij} \tag{9}$$

while the average response time over all market place images is calculated as

$$\bar{t}_{\cdot j} = \frac{1}{9}\sum_{i=1}^{9} t_{ij} \tag{10}$$

The total average $\bar{t}_{\cdot\cdot}$ over both factors would therefore be

$$\bar{t}_{\cdot\cdot} = \frac{1}{45}\sum_{j=B}^{F} \sum_{i=1}^{9} t_{ij} \tag{11}$$

The interaction of the two factors interview location (α) and market place picture (β) of all locations and pictures is as follows:

$$(\alpha\beta)_{ij} = t_{ij} - \bar{t}_{i\cdot} - \bar{t}_{\cdot j} + \bar{t}_{\cdot\cdot} + \varepsilon_{ij} \tag{12}$$

where ε describes a general error or noise.

For this, the average response time per interview location $\bar{t}_{\cdot j}$ and the average response time per photograph $\bar{t}_{i\cdot}$ were subtracted from the data t_{ij} and the total average $\bar{t}_{\cdot\cdot}$ was added again (see equation 12) to highlight deviations for a given interview location and image. The resulting response times $(\alpha\beta)_{ij}$ (without the error ε_{ij}) are shown in Figure 45.

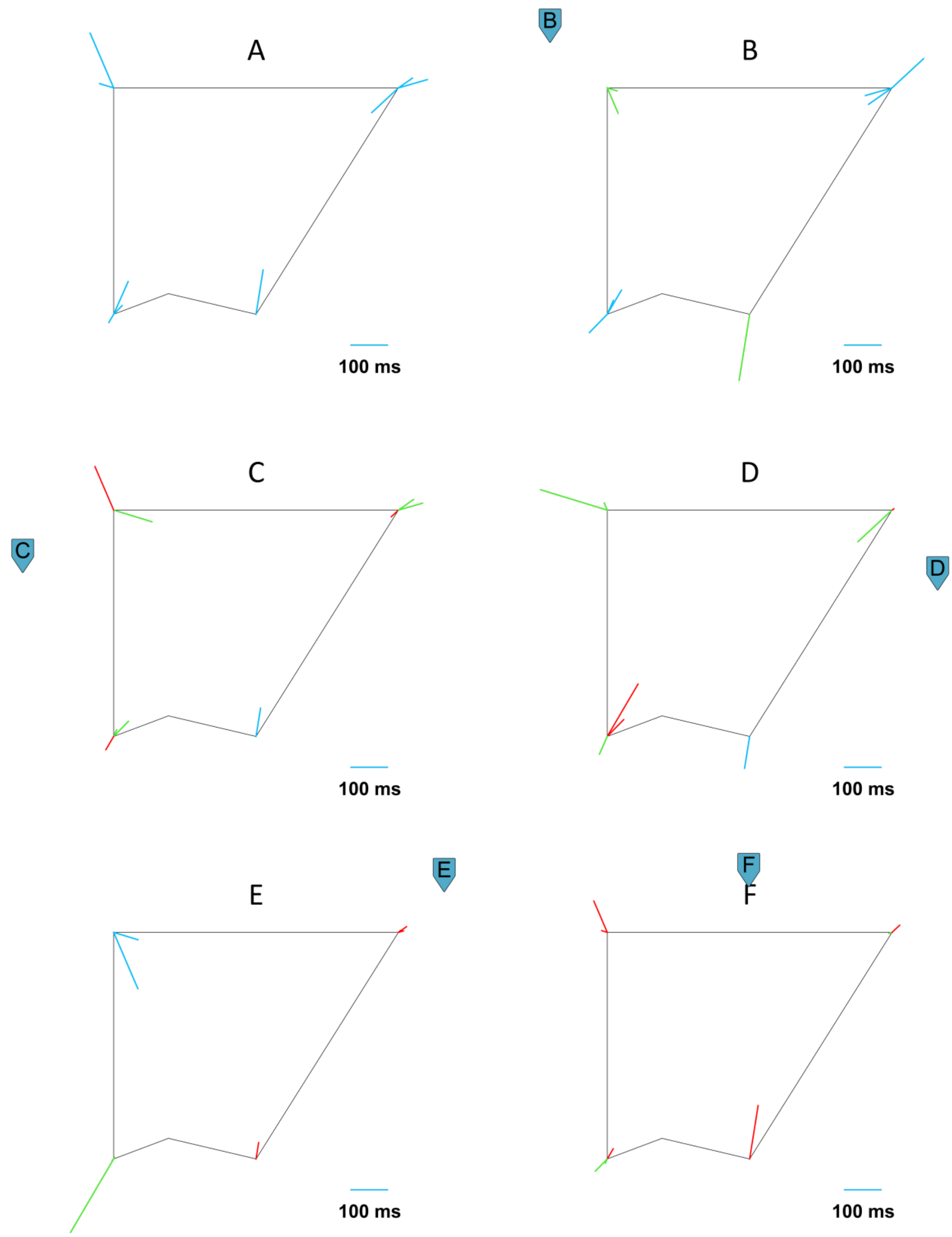

Figure 45. Response time differences after subtraction. Response times varied depending on interview location, view point and angle. Lines pointing inwards indicate shorter response times, lines pointing outwards show responses that took longer than average. There is a tendency that photographs taken at avenues leading to the target place (dark outline, north up) from the respective interview location (blue markers, B - F) resulted in faster response times (green inwards). Green lines pointing outwards indicate responses that were slower but in line with the hypothesis. Response times, which were contrary to the hypothesis, are marked with red lines. Blue lines indicate response times that may or may not be in line with the hypothesis and depend on routes that are not typical, yet feasible for the participants (see discussion). All lines indicate view point and angle of presented photographs of the target place and their length response time differences.

Some tendencies became visible after subtraction (Figure 45). Responses that were faster than average are indicated by lines pointing inwards. Responses that were slower than average are indicated by lines pointing outwards. Green lines indicate responses that are in line with the hypothesis, that is, pictures which show the view of a place that would be seen upon walking directly from the interview location to this place would be recognized faster, while other pictures - especially on the opposite side - would be recognized slower. Red lines indicate responses that are contrary to this hypothesis. Blue lines are equivocal, because depending on the taken route the response may or may not be in line with the hypothesis, but since it is unknown which route the participants would have taken, it could not be determined.

At interview location B, pictures taken at the upper left and some of the lower left entrance were recognized faster (lines pointing inwards). Also at the upper right entrance, two of three pictures were recognized faster. In contrast, pictures taken at the lower right entrance as well as one at the lower left and upper right entrance were recognized slower (lines pointing outwards). This is mostly in coherence with the direction of this interview location (indicated by the blue tag in the figure) relative to the location of the target place. Looking at the map (Figure 37), typical paths from B to the target place are either first eastward and then southward, or first southward and then (south-) eastward. Participants are not supposed to enter the market place from the lower right side. The upper right avenue could be used but is not the shortest path.

For interview location C, one of the upper left and two of the lower left (green inwards) as well as the lower right (blue inwards) pictures were recognized faster, which is in coherence with the location of C to the target place. Also, the two pictures in the upper right corner, which were recognized slower (green outwards), support the theory. However, there was one picture both at the upper and lower left that was recognized slower (red outwards), and one at the upper right entrance that was recognized faster (red inwards), which are contrary to the hypothesis.

At interview location D, two pictures taken at the upper left and one at the lower left (green outwards) as well as in the lower right corner (blue outwards) were recognized slower, supporting the hypothesis. One picture at the upper right was recognized faster (green inwards), which is also in line with it. However, two pictures taken at the lower left entrance were recognized faster, too, which is against the hypothesis.

Participants asked at interview location E recognized pictures taken at the upper left entrance (blue inwards) faster, while the majority of those taken at the lower left (green outwards) were recognized slower. Still, the picture on the lower right (red inwards) entrance was recognized faster, too. Also in the upper right, the images taken there were recognized slower.

For interview location F, two pictures (one lower left and one lower right, red inwards) were recognized faster, however, this was contradictory to the hypothesis. All the other pictures were recognized slower (red outwards), where only the two in the lower left corner (green outwards) are in line with the hypothesis. Altogether, this interview location showed the poorest results in regard to the hypothesis.

At interview location A there was no reasonable path leading to the target square as it was several kilometers away. Therefore, it is not evaluated and just added for the sake of completeness. Also, interview location M (directly at the market place) was not evaluated for that reason [see appendix Figure 61 b)].

These findings show mixed results: Some response times support the hypothesis, some are contrary to it. This is why in the next step a statistical approach was chosen.

For each near interview location (B, C, D, E, F), the oriented response time differences (red, blue and green lines in Figure 45) were treated as vectors. These vectors (black arrows, Figure 46) depend on the response time differences and are shown with their angular deviation from the air-line distance from each interview location towards the point where each photo was taken and its orientation (green line, Figure 46), similar to the vector plots of experiment 4. In Figure 46, the 45 response time vectors (9 photographs x 5 interview locations) of all near interview locations (B to F) averaged over participants are shown. The longer and the closer the black arrows were to the green line, the stronger was the effect of oriented views as it was expected in the hypothesis. However, the vectors point in all directions, indicating no clear preferential direction and therefore no general effect for the individual pictures. For the vector plots of each interview location see appendix Figure 62.

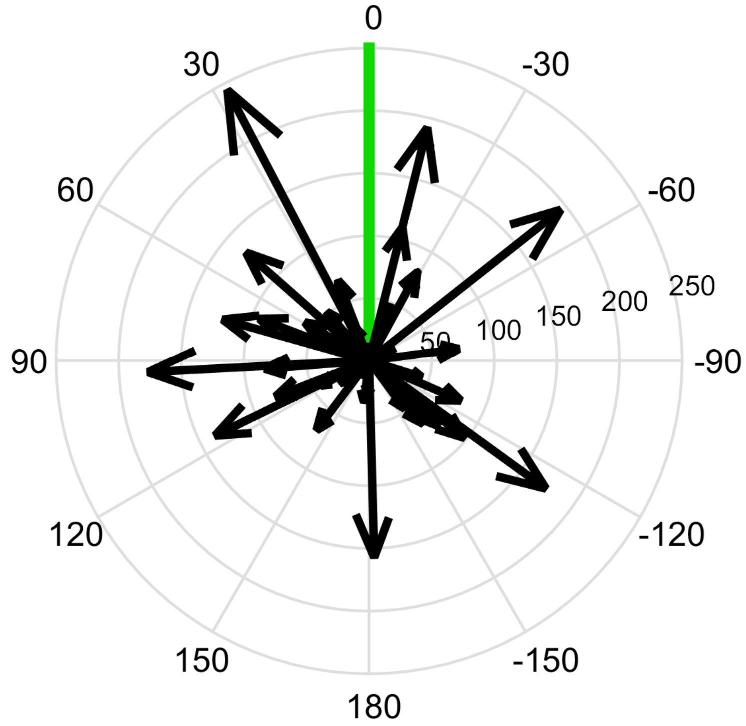

Figure 46. Response time vectors, rotated to align the air-line directions (green line) from the near interview locations (B to F) to 0 degrees. Vectors (black arrows) are randomly distributed and do not show a directional bias. Radial numbers display response time differences in ms.

In the final step of analysis, pictures were grouped and response times analyzed according to the four avenues and entrances (Figure 47) leading to the market place. A two-way repeated measures ANOVA of interview location and entrance reported significant main effects of interview location, $F(6, 132) = 3.418$; $p < 0.01$, and also of the four entrances, $F(2.50, 330.15) = 4.729$; $p < 0.01$. No interaction between the two factors was present. Since Mauchly's test indicated that the assumption of sphericity had been violated for entrance, Greenhouse-Geisser corrected values were used here. Bonferroni corrected post-hoc tests showed significant differences ($p < 0.05$) between NW and SE at interview location A. At interview location D there was a significant difference ($p < 0.05$) between NW and SW (all significant differences are displayed with stars in Figure 47). No differences at the other interview locations could be found.

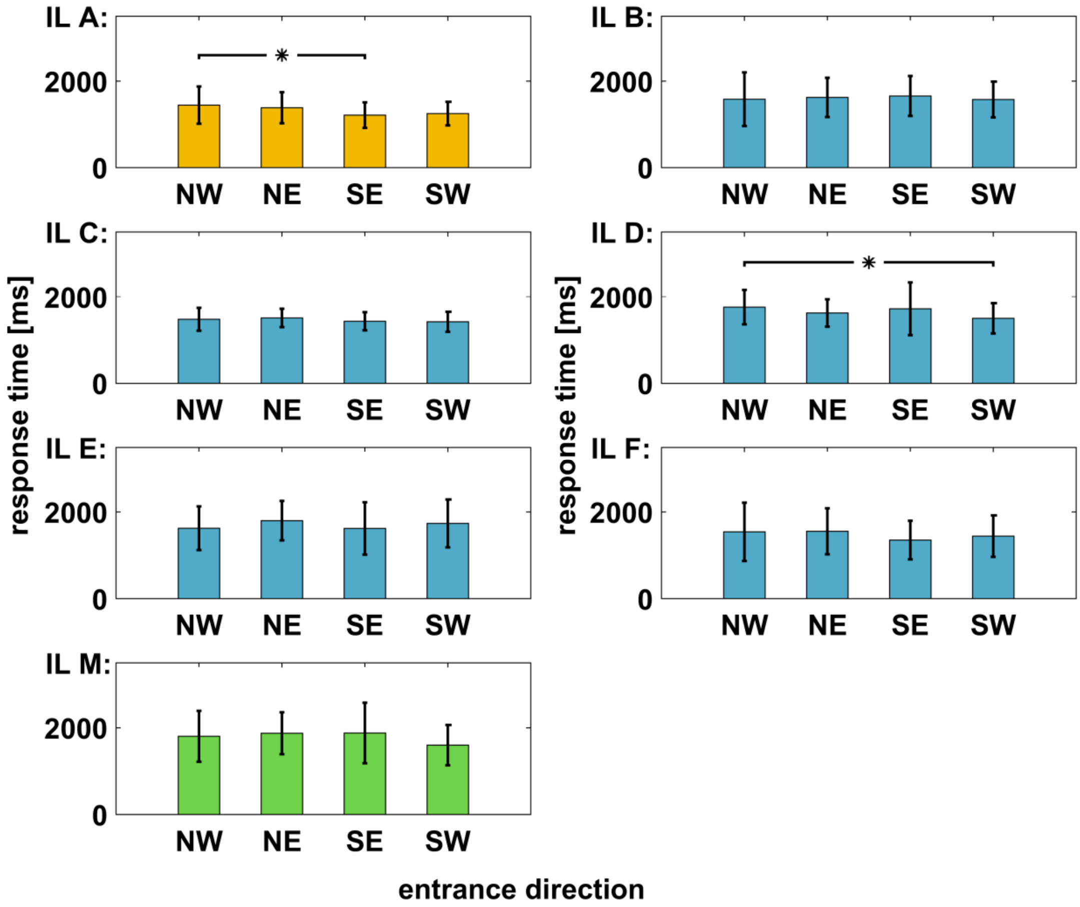

Figure 47. Average response times per entrance to target place for target place photographs. Varying response times depending on interview location and on target place entryway (NW, NE, SE, SW) were present. The y-axes show the response times in milliseconds, the x-axes the entrance directions. Error bars show the standard deviation.

Taken as a whole and summarizing the results of experiment 5, pictures taken at the potential entrances from which a participant is likely to enter the target place from his current position are recognized faster. However, there are also pictures from other entrances which are unlikely to be used and still response times were small.

Participants were asked for their estimation of the distance between the interview location and the target place in minutes' walk on a questionnaire. No correlation between the estimated walking distance and the response times [Figure 48 a)] could be found. However, a correlation between the veridical distance (air-line distance measured on a map) and the response times was present, $r(118) = -0.230$, $p < 0.05$, with a negative slope of -1.22 [Figure 48 b)]. Participants' response times decreased with increasing distance from the target place.
To evaluate the participants' performance in estimating the distance in minutes' walk, Google Maps' (Google LLC, USA) pedestrian navigation was used to get "typical" durations for walking from the respective interview location to the next nearby entrance to the target place. Participants' estimations were significantly longer in three out of five cases as t-tests against Google's estimations (black dashed lines) revealed (Figure 49). Interview location B: $t(19) = 2.476$; $p < 0.05$. Interview location E: $t(19) = 3.214$; $p < 0.01$. Interview location F: $t(19) = 3.312$; $p < 0.01$. Also in Figure 48 a), it can be seen that the participants' estimations were varying considerably, especially for close interview locations like F and B.

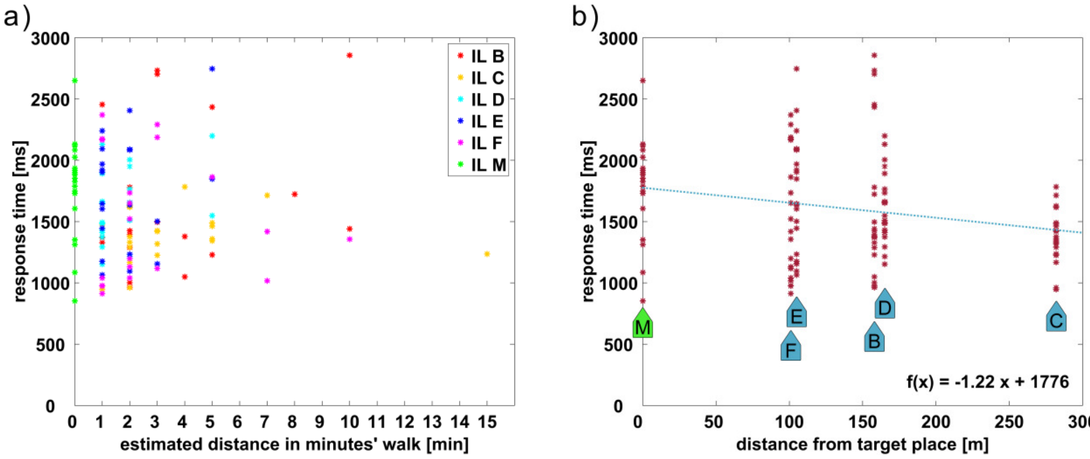

Figure 48. Correlations between distances of interview locations and response times. a) No correlation between the estimated distances in minutes' walk and the response times could be found. Colored dots mark single data points (one participant each) at the respective interview location (IL). The y-axis shows response time, the x-axis shows the estimated distance from the interview location to the target square in minutes' walk. b) Between the air-line distance and the response times, a correlation with negative slope was present. Red dots mark single data points; the blue line indicates the regression line. Letters indicate interview locations. The y-axis shows response time, the x-axis shows the air-line distance from the interview location to the target square.

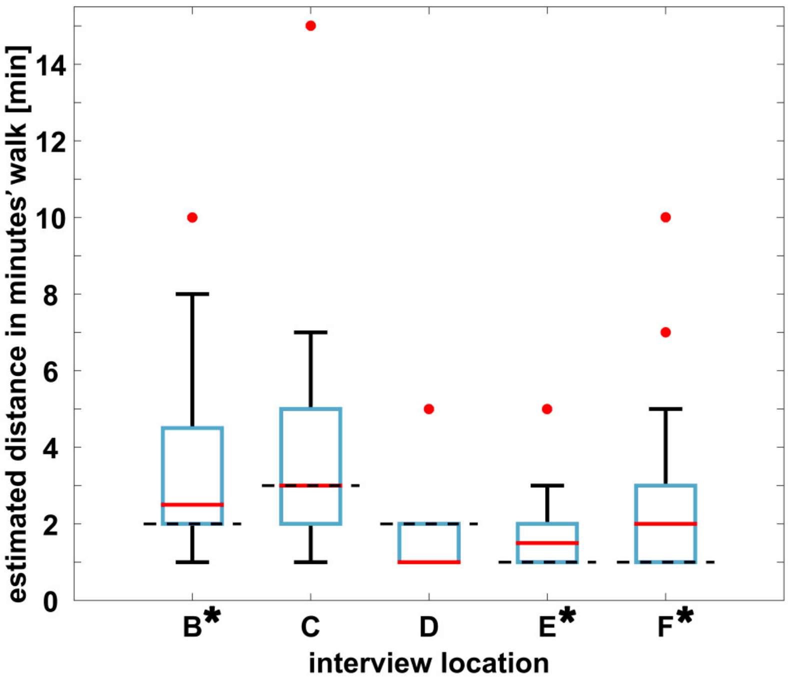

Figure 49. Comparison of estimated distances in minutes' walk. Participants' average estimations of distance in minutes' walk from the interview location to the target place were significantly longer than Google's estimations (black dashed lines) for interview locations B, E and F (marked with a star). Red bars show median of participants' estimations, blue boxes show areas between 25 % and 75 % quartile, black error bars show maximum and minimum data within 1.5 times interquartile range, red dots show data outside of that range. The y-axis shows estimated distance in minutes' walk, the x-axis shows interview location.

9.3. Discussion

Passers-by were interviewed at different locations in the city of Tübingen. They were shown pictures of different places on a tablet computer and asked to decide whether these pictures show a certain place (market place of Tübingen). It was hypothesized that pictures that the participants would see first when entering the target place from their current location would be recognized faster.

Participants were performing well and correctly identified most of the presented pictures with a d' of 3.8. This shows that participants were familiar with the views of the selected places in the city of Tübingen and most likely have spatial knowledge about their relationships.

No gender effect for response times could be found. Both male and female participants had about the same response times on average for solving the task. This seems to be contrary to the prevalent notion that males typically outperform females in certain spatial tasks (McGee, 1979), like wayfinding, navigating (virtual) mazes and mental rotation (Moffat et al., 1998; Malinowski & Gillespie, 2001; Astur et al., 2004; Andreano & Cahill, 2009); which might be related to the right anterior hippocampus (Wei et al, 2016). As a side note, these superior spatial abilities can be caused by early exposure to testosterone during child development (e.g. Hooven et al., 2004), and this advantage also can be transferred to females via intra-uterine hormone transfer (Heil et al., 2011).
On the other hand, it has been shown in several studies that female participants report more landmarks in spatial and route-learning tasks than male ones (McGuiness & Sparks, 1983; Miller & Santoni, 1986; Galea & Kimura, 1992). The task in experiment 5 was to recall and recognize (familiar) views of places, which could also slightly favor females. In the end, both advantages probably led to equal response times in total.

Response times varied with respect to the different interview locations. Participants were slowest on average if they were interviewed directly at the target place (market place). It is likely that participants were doing a slightly different task here, and different brain processes were active. Participants were - in addition to the main task - probably also (mentally) taking the perspective of the currently presented picture of the target place and their current location, although they were not instructed to do so. This of course needs additional processing resources of the brain and would result in increased response times. Evidence for this process comes from Wiener & Mallot (2003) who found that there is a fine-to-coarse process, at least in route planning in regionalized environments, depending on the distance from the target location. The same process could be present here as both the target place and interview location in this special case can be seen as "target region" compared to the other interview locations (also see below).

An interaction between nearby interview locations and pictures of the target place and also other places could be found when they were grouped into cardinal directions. This means that certain pictures of the target place were identified faster or slower depending on the interview location, indicating that pictures that show views which are likely to be seen when approaching the target place are identified more easily, which supports the hypothesis. However, this interaction could not be found for individual photographs (see below) but only accumulated into cardinal directions.
This is also well illustrated in the vector plots: While some arrows align with the air-line direction between interview location and target place, other arrows do not and even go into the opposite direction. This is dependent both on the images and interview locations as can be seen in Figure 62. In fact, at interview location "F" the poorest results were obtained (also see Figure 45). Yet, it is not known if the location "F" maybe is not suitable for that kind of task, e.g. because it is directly between two avenues that lead to the target place, and was selected in an unfortunate

way, or if participants there differed from those at the other interview locations, e.g. following other (navigational) intentions, or if it was just accidental.

A detailed analysis of the nine photographs taken at the different entrances leading to the target place indicated a tendency congruent with the hypothesis, which means that pictures that one would see upon entering the target place were recognized faster. Though, this could not be found in all cases and for all images and entrances. But then, it is difficult to assess which routes the participants would have used to approach the target place from their interview locations without giving cues. Response times are dependent on the participants' typical routes, though, because these routes lead to the respective views that in turn may be recognized faster. This is why there are certain cases and routes, (Figure 45, blue lines) that do not correspond to the shortest route, yet they are still feasible to be used by the participants without being a longer detour. Also, directly asking them after the experiment could again lead to different results as oral descriptions of routes tend towards "explaining the way" in which the explaining person, among other things, makes assumptions about the map knowledge of the other person. This means the participant would not essentially describe the route he or she would have used on their own, or they focus on different aspects and landmarks while explaining (Denis, 1997; Lovelace et al., 1999; Hölscher et al., 2011).

Some of the photographs taken at the four entrances and used as stimuli for the market place were maybe chosen poorly for the task. Photographs varied in field of view and viewing angle depending on the entrance. Care has been taken to display sufficient information of the surrounding with the images; however, this could deviate from the internal representation or memorized view of the participants, leading to larger response times than normal. It is also possible that different views at a given entrance are active in memory depending on the approach direction (and hence interview location), which would lead to seemingly inconsistent response times at that entrance. For example, a participant approaching the market place from interview location C and entering it at the north-western entrance typically turns left immediately in his everyday route. This would result in small response times to one of the images taken at that entrance but not to the other image since that one points into a different direction and shows a different angle. In fact, this result is displayed in Figure 45 for that interview location. While both images taken at that entrance are meant to display typical views when entering the market place at that point, they could have different meanings to the participants in their wayfinding or route planning and therefore are associated differently in their view graph. This would also explain why an interaction between interview location and accumulated pictures of cardinal directions could be found but not for individual pictures.

The analysis of the response times in respect of the four entrances themselves showed an effect, indicating that certain entrances are more prominent than others. However, no interaction between entrance and interview location could be found. An explanation for this is that there are no fixed routes from a given interview location to a certain entrance, but that there are several feasible options to approach and enter the market place.

In the experiment, a negative correlation between participants' response times and the air-line distances between the nearby interview locations and the target place could be found. The fastest responses were found at the distant interview location. They were not considered for the correlation analysis but are in line with it: The further away the interview location was, the faster the participants responded. This is also in accordance with the results found in experiment 4 and gets support from Wiener & Mallot (2003) and their "fine-to-coarse" navigation model. However, there was no correlation for the estimated (walking) distances. This could be due to disparities between air-line distances and walking routes. While air-line distances show the direct positional relation, walking routes also cover turns and detours because of the geometry and layout of the town. Nevertheless, in this downtown setting, actual walking routes here are only marginally longer than air-line distances (30 m in the worst case) as both measurements by hand on a map and Google Maps' pedestrian navigation pointed out. According to the comparison of estimated distances (Figure 49), it is very likely that the lack of correlation here is caused by poor distance and walking speed estimations of the participants for that region. Yet, it has to be noted, the quality of Google's speed estimates of people walking in that specific area is unknown, too.

This experiment was carried out in the same inner-city area like experiment 4. Therefore, the same conditions and limitations apply, namely the architectural and city planning layout of the buildings and roads as well as the geographical features. There is a slant from north to south and a sharp border (river and city wall) that blocks access from the south to most parts of the inner city. This is the reason why there was no near interview location in the south as it would lead to unpredictable results because it could be attributed to a different region (see discussion of experiment 4). It has been shown that geographical slant can facilitate navigation and orientation (Restat et al., 2004). This also could have assisted the participants solving the task in this experiment, although, slant was not always visible on the photographs. Still, the target place has a noticeable slant which could help to rule out at least some of the non-target pictures. These given conditions in the inner city could also explain the smaller response times for all pictures which show scenes in eastward and southward directions (Figure 43).

All these features, among others, lead to regionalization of the environment, which in turn helps to navigate and orientate. Furthermore, a regionalization could explain the negative correlation of response times and distance that has been found: The further away, the faster the participants responded. According to the "fine-to-coarse" route planning theory of Wiener & Mallot (2003), distant locations are represented with coarse space information compared to nearby ones. Since less detailed information needs to be considered at distant interview locations in this task, it would allow for smaller response times. This is also in line with the largest response times being found directly at the target place.

Altogether, the results and tendencies found here are in line with the results of experiment 4 and give new insights to the view-based model. It gets support from the data of this experiment; however, some results are ambiguous. It has been found that certain views were recognized faster in the direction of approach as it was hypothesized. Admittedly, there were also responses that were contrary to the hypothesis. The stored and retrieved "views" from memory seem to be more complex than just "simple photographs" that are stored and looked up according to the direction of (mental) approach in a photo album.

10. General discussion

Five experiments were developed and carried out in this thesis to investigate the representation of faces and places and the influence of the actor on these. Experiments 1, 2 and 3 focused on face perception and representation in regard to different aspects: Face recognition under different conditions and tasks (categorization, discrimination, identification, primed and unprimed) as well as face characterization (similarity and distinctiveness) were investigated.

In experiment 1, participants were asked to solve discrimination and categorization tasks with morphed pictures of faces of themselves and other people. It was found that with the given faces and familiarity level, participants did not exhibit an increased performance in a discrimination task when crossing the categorical perception boundary - for both self and other faces. Beale & Keil (1995) found an increase of discrimination performance upon crossing the boundary of morph sequences of famous people; however, they did not find such an increase for faces of unknown people. Also Bülthoff & Newell (2004) did not find such an increase for sex judgements of people's faces unless artificial sex continua independent of identity were used and participants were trained thoroughly. It could be that the participants who took part in experiment 1 of this thesis were lacking familiarity with the faces of the other people. However, they were supposed to know their own face well enough and therefore should be able to at least discriminate between a familiar "self" and some unfamiliar "other face".
Furthermore, it was found that the variance in the categorization task was larger for morph sequences that included faces of the participants. This means that the PSE-level was also varying to a larger degree, and therefore the categorical boundary between two faces was less uniform for the used stimuli. This lack of a sharp categorical boundary can also explain why no peak in discrimination performance could be found in the data - as it was hypothesized and found by Beale & Keil (1995) in their study. If the boundary is indistinct and "blurry", so is a potential performance increase, at least for "self" faces. This could also give a rationale to the results of Bülthoff & Newell (2004). It is conceivable that for morphed images between sexes the categorical boundary (for sex judgments) is also strongly dependent on which images were actually paired and that there is not a sharp (general) boundary, too. Unless participants were thoroughly trained to spot individual characteristics of each face in this regard, no increased performance at the categorical boundary could be elicited.
Nevertheless, in experiment 1 of this thesis, an increased discrimination rate for pictures with "self" was found. Furthermore, if "self" was present in the morph sequence for 50 % or more, an increased performance can be assumed. This supports the idea that participants had a well-learned mental image of "themselves". It also points to a stronger or at least different representation of "self" compared to other persons.

Experiment 3 focused on the perception of identity and how it changes with periliminal primes. Therefore, participants were presented with brief presentations of faces taken from different morph sequences that included themselves and other faces. Participants had to make judgements about the identity of the presented faces. In both experimental setups [3 a) "self", 3 b) "other with text"] as well as in experiment 1 without priming, the response time curves were combinations of a linear and a gauss function with different slopes. Near the 50 % level, responses were slower since ambiguity of the pictures was largest. The more unequal the mixing ratio between the morph pictures was, that is towards the 100:0 % and 0:100 % levels, the faster the participants' responses were. In the case of priming, these small response times were modulated by the congruency between prime and test picture [experiment 3 a)]. Furthermore, if text was used as instruction [experiment 3 b)], it already interacted with the prime and test picture and led

to flat response time curves at the lateral ends - similar to the response time curves of experiment 1. In cases where the person in the instruction matched the prime picture [experiment 3 b)], response time curves were again similar to those in experiment 3 a), since a congruent priming effect was taking place. Congruency here means congruency between instruction and prime image. This priming effect led to shifts of perceived identity, which became apparent in shifted PSE-levels. These shifts were in line with the respective prime picture in all experiments. Still, the shifts were smaller in experiment 3 b) with text as instruction.
To explain these results, a model was designed that can give account to these findings. Apparently, the instruction to think of a certain person (who is different from oneself) causes to activate a representation of said person, which then - in accordance to the IAC model - activates connected nodes. This pre-activation leads to an (additional) facilitation of the response once the matching (prime) image gets displayed. However, if these two stimuli do not match, confusion takes place, which cancels out priming effects on the test picture. This is reflected in both response time curves and in weaker shifts of perceived identity. It has to be noted, especially in regard to the studies by Bruce & Valentine (1986) and Ellis et al. (1987), that the categorization experiments in this thesis were requiring at least some representation of the presented faces and their identity, particularly in experiment 3 b) where participants had to judge between other people, or else they would not have been able to solve the task. In the two mentioned studies, however, it was not necessary to use the text or names to solve the task.
The different magnitudes of the shifts of perceived identity in experiment 3 a) and b) point out to a different representation of “self” in comparison to other people. Participants were instructed to identify the person in the shown picture as oneself and press the according button on the keyboard [yes-no-paradigm; experiment 3 a)]. Experiment 3b) was also set up in a yes-no-paradigm for reasons of comparison. Here, participants should identify the shown picture as target person and press the according button. Instructions were fairly comparable considering the experimental setups and should not favor a particular representation “from the outside”. Admittedly, in experiment 3 a), it was always asked for the participants themselves, and therefore the target person stayed the same, while in experiment 3 b) the target person was changing with the names. This could have led to a weaker representation, which could explain to some degree the smaller shifts in perceived identity in experiment 3 b). On the contrary, a weaker representation from the instruction would also have had less influence on the response time curves, and therefore the significant priming effect - as it was present in experiment 3 a) - should not have been canceled out completely in experiment 3 b) when instruction and prime were not matching. That is why it is concluded that the representations of the target persons were strong enough to cause significant shifts in perceived identity and interact with the visually presented prime picture. Nevertheless, the representation of oneself was still stronger than the representations of other faces as the magnitudes of the shifts were showing.
The results of these experiments are interesting as they additionally give insight into priming processes. With the progression from experiment 3 a) to 3 b), and along with the studies by Ellis et al. (1987) and Bruce & Valentine (1986) as well as Calder & Young (1996), it could be shown that for a priming process to take place, it is not only important which stimuli are present but also their relevance for the task. This is illustrated in the presented priming model.

Various studies were investigating which brain regions are involved for recognizing faces and especially familiar faces. As stated before (see introduction for experiments 1 to 3), a region in the fusiform gyrus in the temporal lobe, named fusiform face area, was identified and named responsible for recognition of faces and face-like objects (e.g. the famous yellow smiley face), while the occipital face area especially focusses on certain features of a face (e.g. mouth, eyes …). However, it is debated if the fusiform face area contributes to other expertise recognition tasks, too, like identifying certain car models or bird species (Gauthier et al., 2000).
Studies show that different familiar faces elicit different activation patterns in the fusiform face area (Nestor et al., 2011; Ghuman et al., 2014), as well as in the anterior fusiform gyrus, indicating

a role in person identification. Kircher et al. (2001) found activations in the right hippocampus, insula and anterior cingulate cortex and left prefrontal and superior temporal cortex upon contrasting images of "self" with unfamiliar faces. Once they compared images of the participants themselves and their respective partner (used here as stimulus of a very well-known face and important person), only the right insula region showed increased activation, indicating this as a resource for "self" processing. Leveroni et al. (2002) found activations in several areas of the prefrontal cortex and temporal regions as well as hippocampus and parahippocampus when participants viewed famous faces compared to newly learned faces. In contrast, the newly learned faces led to stronger activations in the frontal and parietal regions of the brain. Leube et al. (2003) also found stronger activations for newly learned faces in the left inferior parietal and medial frontal cingulate cortex. Kircher et al. (2001) compared brain activation levels and regions while showing pictures of the participant, their respective partner (both well-learned faces) and unknown faces. Increased blood oxygenation in the right limbic system, superior temporal cortex and left prefrontal cortex when viewing self-faces in contrast to unknown faces was present. An increased activation of the right insula was found comparing partner vs. unknown images. They conclude that "the combination of right limbic and left cortical activation could underlie human self-recognition" (Kircher et al., 2001). Also Taylor et al. (2009) contrasted images of faces with different familiarization levels with each other. All familiar faces showed activations of the fusiform gyrus, while the faces of the participants ("self") also activated occipital and parietal regions. The activation of these different brain regions is also expected to take place in experiments 1 and 3 a), at least, where faces of the participants themselves were compared with (learned) faces of other people. Of course, functional brain imaging is required to know exactly which areas are active in these tasks. Faces of the respective partner additionally activated the parahippocampus and middle temporal and frontal gyri. The cingulate cortex was activated when personally familiar faces were contrasted with unfamiliar ones. Lucas et al. (2003) investigated the lateralization of face recognition. While the right temporal lobe was responding "earlier, and with uniform frequency reductions" compared to the left one, they concluded that not only the anatomic location in the brain is an important factor but also the timing and activation spread through "widely distributed neural ensembles". Keenan et al. (2000) were also investigating the location of self-recognition in the brain. They say that the prefrontal cortex could be a preferential component for this task, with a potential bias for the right side within this structure. However, they also stated that it is "highly improbable that there is a 'self-recognition' or 'self' center". For an overview of the involved brain regions and studies in which they were investigated see Devue & Brédart (2011).

In summary, there are many different brain regions involved in the recognition and processing of faces. In early stages of processing, like in the occipital face area and fusiform face area, the detection of faces and face-features in a visual stimulus is focused on. Other brain areas are active if it comes to identification of the viewed face and whether it is familiar, famous, newly learned or known in a different way. Depending on the familiarity, associated brain areas get activated, too; e.g. long-term memory brings past episodes where this person was met, the amygdala contributes which feelings this person elicits, etc. Furthermore, in an experimental setting, these activations also depend on the task the participant has to complete. Experiments often differ between passive viewing of images and solving a task like making familiarity judgments or the like, which might lead to different activations of (further) brain regions.

In experiment 3 a) of this thesis, participants had a well-known face as target face, namely their own face. In experiment 3 b), in turn, they were shown different target faces they should have familiarized themselves with earlier. Also in experiment 2, participants were given printed cards with faces on them to memorize these. This process that turns a formerly unknown face into a familiar one and what brain regions are active at different levels of familiarity (see above) were investigated in different studies. Also, how familiar faces can be detected under all kinds of

lighting and viewing conditions and even if facial parameters have changed (different hair style, glasses, make-up …) was focused on in research.
It is believed that regular exposure and if possible interaction with a certain face (and therefore the person) or an ethnical group of faces will adjust the face-space so that it can detect variances in this subset of (face) stimuli better. In both theories of face-space - with and without a central prototype face - the space can have different local densities and resolutions. If a certain group of people are seen and interacted with regularly, the brain readjusts that region in the face-space to which the faces of these people belong to, so that they can be distinguished easily from another and also from other faces one has met (see also below).
For a single face, it is learned that a greater variety of visual inputs (lighting, hairstyle, …) still refers to the same person. A recent study by Kramer et al. (2018) was modeling familiarity in a computational approach with thousands of images. They concluded that becoming familiar with a face is an "increasingly robust statistical description […] of within-person variability". This knowledge about image variety is thought to be "programmed" into a face recognition unit in the IAC-model, which means the face recognition unit is tuned more precisely to that face and responds less if a similar yet different face is shown. However, this ability seems to be located elsewhere and not in the early face processing regions like fusiform face area, occipital face area or posterior temporal sulcus, as a study by Davies-Thompson et al. (2012) revealed; at least they did not find different levels of activation for image-invariance between familiar and unfamiliar faces in these areas.

With these findings and new technology like functional imaging, an alternative model of face recognition has been proposed by Haxby et al. in the year 2000. It shares some elements with the (extended) IAC-model by Burton, Bruce and Young (see above), but focuses on functional anatomical differences that have been found and identified with imaging and lesion studies. In their model, Haxby et al. differentiate between a "core system" in which visual face recognition and perception takes place, and an "extended system" that is recruited to support the core system in its function and to add additional information. The core system consists of the inferior occipital gyri, that provide an early perception of facial features and forward it, the superior temporal sulcus and the lateral fusiform gyrus.
The lateral fusiform gyrus, which also includes the fusiform face area, is thought to activate on and detect invariant aspects of faces; that means, the identity of a person [although Davies-Thompson et al. (2012) did not find identity-based activations there, see above]. It is both linked to the superior temporal sulcus of the core system and to anterior temporal regions, which are part of the extended system and supply with information on personal identity, name, biographic information, events etc.
The superior temporal sulcus of the core system, however, is recruited to detect changeable aspects of faces like expressions, movements of the lips and face muscles, and eye gaze. This information is used for interaction with the person and to detect the emotional state, intentions, current focus of attention etc. To accomplish this, it gets support from the extended system, namely the intraparietal sulcus for spatially directed attention, the auditory cortex for prelexical speech perception, and the amygdala, insula and limbic system for emotions.
The extended system is not exclusive for face recognition and perception but also used to fulfill other roles, e.g. to control one's own emotions or for spatial attention on any object. However, its specialization can be recruited to aid the core system for face perception.
Haxby et al.'s model is not contradictory to the IAC-model but rather extends it. In the IAC-model there was no differentiation between invariant and changeable aspects of faces. Also, it does not mention involved brain structures but is solely a functional description. Nonetheless, it is still valid to explain separate functions in face perception and how priming can influence the processing. Both models are useful to understand face recognition and representation, especially in connection with the identity of a person and knowledge about the person. They also help to explain the processes of identity priming like it was used in the experiments of this thesis.

Still, an adjusting face-space is a core theory in face representation. In experiment 2, a small segment of the face-space with 2 x 5 faces was investigated. It is debated if there is a central prototype from which different faces are encoded as distances in feature space (one might say an “allocentric encoding” in analogy to spatial cognition) or if only the differences between faces are encoded, which span this multi-dimensional feature space without a central “anchor”. Valentine et al. stated in 2016 that “the central tendency of the relevant population is defined as the origin for each dimension. Thus the density of faces (exemplar density) is greatest at the origin of the space. As the distance from the origin increases, the exemplar density of faces decreases. The faces near the origin are typical in appearance.” This also means a prototype could be inferred from the face-space, but it is not necessary to create that space. Furthermore, they described the distribution of faces in the face-space: The closer to the “origin for each dimension”, the more faces are represented here.

The recognition of faces in this face-space, and also the space itself, was addressed by Lewis & Johnston (1999). They suggested a multi-dimensional Voronoi diagram for faces (for an example of a Voronoi diagram see appendix Figure 63), which separates the space into “cells”. Inside this cell is the veridical face. The borders around the cell separate the face equidistantly from all neighboring ones, while the area inside the cell describes “faces” and features that are closer to no other familiar face. This can be seen as a “capture area” for face recognition where a presented picture by perception would fall in a certain cell and then be matched to the face which “defines” that cell.

In the face-space, faces which are more typical would be placed closer to the center or origin of this space. More distinct faces would be found further away from this center. Such distinct faces were found to be recognized and remembered more easily (e.g. Light et al., 1979). In experiment 2 of this thesis, the distinctiveness of and similarity between the faces used here were investigated to characterize these faces. Furthermore, it was examined whether there were systematic shifts in perception between certain face pairs, and finally if there was a correlation with the face characteristics. Only marginal shifts for very few faces in general were found. However, the number of individual shifts in the face-space was substantial but not uniform. This shows that the face-spaces for these faces used in the experiment were different for each participant, and shifts were dependent on the individual and respective face pair. Male faces were categorized more often as “distinct” compared to female faces, which is in line with other studies (e.g. Enlow, 1990). No correlation between distinctiveness and shifts could be found for female faces, while a tendency was present for male faces, indicating shifts towards the more distinctive face. However, for female faces there was a correlation between shifts and similarity between faces: The stronger the similarity, the larger the shift. An explanation for this unilateral effect could be ambiguity. Female faces were rated less distinctive, which means they could also be confused more easily, especially if the differences between neighboring morph-levels were low. This would result in (larger) shifts - which exactly was the case in the experiment. In the face-space model, these faces would be closer together with only small differences in the feature space. According to the model, these small differences - and therefore large ambiguity - between certain faces could be overcome by regular exposition and interaction with these persons, because then the sensitivity for the feature differences would increase and a (local) change of scale would be introduced in that face-space region. In the Voronoi diagram, it would mean to enlarge (certain areas within) cells that are small or in which the veridical face is close to one or several borders by distorting space so that the cells cover a larger area and the veridical face is located more centrally, visually speaking.

In experiments 1, 2 and 3, the representation of faces and the role of “self” was investigated. These representations are not rigid but flexible in several ways. A long-term effect is achieved by repetitive exposure and interaction with faces. This readjusts the face-space and adapts it to the faces that surround the individual. A short-term effect is achieved by priming, which alters the perceived identity of a person, both of oneself and of other persons. An effect of intermediate duration is achieved by memorizing faces, e.g. in learning tasks. The priming effect of text/names on face recognition depends on the task. If names are required to solve the task, they are also prone to priming. The own face plays a special role in face representation as it is best memorized and recognized, leading to the strongest shifts in a priming paradigm as it was used in this thesis.

In the second part of this thesis, the representation of places and the interaction with oneself was in focus. In experiments 4 and 5, it was investigated how locations are represented. In experiment 4, passers-by in the city center of Tübingen and at distant locations were asked to sketch two well-known places of downtown Tübingen. Participants' sketches were oriented in dependence of their location while drawing - however, only when they were in the proximity of these places. At the distant locations, produced sketches had a general orientation. In experiment 5, passers-by in the city center and at a distant location were approached and asked to identify the market place on photographs of different places of Tübingen. Results show that images that showed the market place in the direction of approach from the interview location were identified faster. However, also other images were sometimes identified faster although these were unlikely to be used as direction of approach. Furthermore, there was an effect of distance: The further away the participants were, the faster they identified the photographs.

In contrast to faces (and many objects), locations cover an area and typically cannot be referred to as being located at a certain single point because of their expanse. Still, the brain manages to represent them, too, and even enables us to use this information for complex tasks, like planning a route through multiple locations. Yet, the brain faces similar challenges to learn and familiarize a new location so that it can be recognized (quickly) from different viewpoints and viewing conditions like it has to do when learning a new face, for example.

For the immediate environment surrounding an observer, the terms "spatial image" and along with it "spatial updating" have been established to describe the representation in which near objects are stored in memory. This representation is ego-centric and is updated immediately through body movements (turning and walking). However, in long-term memory, familiar environments are stored in a different way because otherwise the consecutive ego-centric updating would consume more and more processing resources, the more items or regions are memorized. With an allocentric representation, in turn, such updating is not necessary because storage happens with a general orientation. Though, it is debated what such an allocentric storage in long-term memory looks like, and also how the required transformations are applied when it gets read out of long-term memory and transferred into (ego-centric) working memory. One way of storing places in long-term memory is the view-graph model by Schölkopf & Mallot (1995). Here, characteristic views of a place are memorized and the collection of these views represents that place. For wayfinding, these views are associated with instructions how to get there (e.g. from another view). For both planning and walking a complex route, one passes through the views and associated instructions to reach the goal (see introduction for experiment 4 and 5).

In experiments 4 and 5, participants were required to recall the representation of a certain place ("Holzmarkt" or "Marktplatz") from long-term memory and transfer it into working memory to fulfill the task, which was either reproducing the representation to some degree (experiment 4) or comparing it to a presented view (experiment 5). All this took place in the proximity of said places or in control conditions far away from it. As results showed there were interactions between the near interview locations and the target places, resulting in oriented reproductions or faster responses. In experiment 4, participants "automatically" oriented their produced sketches of the target places in most cases as if they were walking from the interview location towards that place after a generic orientation bias was removed. In experiment 5, participants often recognized views they would see upon entering the target square faster than other views (see below).

This means the (representation of the) current location of the participant must have interacted with spatial memory recall from long-term memory. This resulted in sketch maps that were oriented in the direction of approach from the interview location, although participants were not instructed to do so, and in faster orientation-dependent response times in experiment 5.
This effect was only present at nearby locations for which instructions how to get from there to the target place were presumably available since oriented sketches were only produced at nearby interview locations. At distant interview locations, only the generic orientation prevailed. Also response times showed an effect of orientation only at nearby interview locations. However, the effect of proximity was not determined by the mere distance but also by the regions and their borders in which the downtown area was divided in. Additionally, the closer the interview location was to the target place in the nearby condition, the slower the participants responded. It is beneficial for a navigator to have a detailed representation of the immediate surroundings and the next navigation step because then he does not miss key features or landmarks that help to keep track of the route. At the same time, working memory and mental processing power can be saved by having only sparse representations of faraway locations or later navigation steps since these are not necessary yet, which can result in smaller response times as it was found in experiment 5. This is in line with findings by Wiener & Mallot (2003) and Wiener et al. (2004) who reported differences in route planning and navigation depending on the regionalization of environments, and it also supports the rationale for other regions as they were hypothesized in experiment 4.

Since it is unknown what these stored views in (long-term) memory look like, experiment 5 was conducted to investigate their nature. In this last experiment, participants were presented with previously taken pictures of places and they had to determine whether the pictures were showing the target place or not. Participants were quite good in distinguishing between target and distractor pictures. The hypothesis stated that views one would see first upon entering the target place while approaching from the interview location would be recognized faster, while other views would be recognized slower. In accordance with the hypothesis there were faster responses for those images when they were grouped into N, E, S, W directions. Also for individual photographs small response times occurred in the direction of approach, though, other images and routes that were unlikely to be used led to fast responses, too. Still, the findings strengthen the idea of recalled views in dependence of one's current location. These views are likely to be composed of more than simple "mental photographs" and presumably also contain e.g. geometric or depth information, which can hardly be conveyed with a photograph on a computer screen. Furthermore, passively viewing probably does not activate the stored memories the same way as actually navigating in the area or planning a route.

The aforementioned view-graph representation of places in long-term memory has several benefits. First, when a new place is learned there is no need for transformations from the (visual) input because the egocentric view can be stored 'as is'. Second, in wayfinding, one can compare the stored views with the current view - not similar to but in a way like snapshots in bees (Cartwright & Collett, 1982) - and quickly determine whether one is still following the correct route. Third, in route planning, one can use these views to mentally travel the desired route and check whether it leads to the desired outcome (goal reached, quick route, scenic route, …). And fourth, when someone has to explain the way to another person, these views can be described to that person along with the instructions how to get there.

In a study by Meilinger et al. (2016), the method of approaching passers-by in different regions for navigational studies was continued and they asked visitors in cafes and pubs around the city center of Tübingen for participation. They were asked to arrange slips of paper with the names of landmarks on a sheet of paper so that they reflect the relations and distances between them and so to say reconstruct a (simple) city map. The authors found that spatial recall depended on the current location since the produced "maps" were again oriented - which is in line with the findings

in experiment 4 and 5 of this thesis. However, the authors concluded that such recall is based on spatial image and reference frame transformations from long-term memory instead of recalled views, because in their opinion such survey knowledge cannot exist as views.
It is known that humans can build up survey knowledge and create a “cognitive map” of their environment if it is visited and traveled often (also see general introduction for experiments 4 and 5). However, in a new and unknown environment, navigation begins with individual routes along landmarks, e.g. from home to a resource (e.g. supermarket) and back. Such routes often consist of landmark identification and action labels, e.g.: “Walk ahead until you reach the church and then turn left.” but also with geometric information like “Walk 50 meters/steps and then turn left until you face the bridge”. The more often a route is traveled, the better it is known (Pearce, 1981). If several routes are learned and memorized, especially when they have crossings with each other, it is possible to advance from this route knowledge to a survey knowledge (e.g. Rossano & Reardon, 1999).
A variety of studies were investigating the acquisition of route and survey knowledge (e.g. Thorndyke & Hayes-Roth, 1982; Golledge et al., 1995; Rossano & Reardon, 1999; Rossano et al., 1999), indicating different advantages and disadvantages of different learning methods and representations depending on the task. In a different study by Meilinger et al. (2015), the authors were investigating differences in spatial knowledge from maps and navigation. They concluded in that study that “map experience before and during navigation is enough to cause the organization of spatial memory along the reference frame rather than along a navigation-based one”, which means that the presence of navigation aids (like a map) can determine the kind in which the spatial memory is formed.

All these studies indicate that there is no fixed or single representation of a (new) environment, and that it depends, among other things, on the task and learning procedures, and which one is dominant at a given stage of familiarity. In fact, in the study by Meilinger et al. (2016) in which they had paper slips arranged, there were several differences in the task and setup compared to the experiments in this thesis:
First, they asked people for participation who were sitting at tables in cafes and bars, in contrast to the passers-by who were recruited on the streets for the experiments of this thesis. The latter ones were probably following a certain route and therefore already executing navigation, which means that asking them to sketch a place was just “entering a new destination” for their current behavior and mode of mental processing. Visitors of cafes, however, were likely in a different state of mind and relaxing or discussing a topic and were maybe not so well aware of their location and/or orientation.
Second, some of the interview locations in Meilinger et al.’s study in 2016 were rather far away or at least out of the down-town region, according to the results found in experiment 4 of this thesis. In experiment 4, it has been shown that places that were too far away or belonging to different regions did not yield oriented views in reproduction. This shows that different representations existed for these places at that time, leading to different behaviors and results.
Third, in Meilinger et al.’s study, the task itself was to recreate a survey map and the configuration of ten predefined places and landmarks - which of course needs to be done with survey knowledge and probably even memorized pictures of a city map, which has likely been viewed at some time by the participants. Also, arranging predefined landmarks is a quite different task than creating a sketch from (long-term) memory. This means that the method of their task could be bringing forth the results and conclusions in a self-fulfilling manner.
In the experiments of this thesis, participants were asked for a single place, which allows them to recall a route from the interview location to that spot. Also in Basten et al. (2012), participants were mentally walking a single route at a time. Therefore, the tasks in this thesis, in Basten et al. (2012), and in Meilinger et al. (2016) were quite different from each other, both in terms of execution and in mental processing. In fact, these experiments and studies do not confound their findings but rather enhance them. As it has been explained in the general introduction for

experiments 4 and 5, there are several levels of representation of spatial information, which are, among other things, dependent on the scale.
Navigation with views can be seen as an intermediate level of navigation and representation, and it fills the gap between the local spatial image, which represents the immediate nearby surrounding up to a few meters, and the large scale map or survey knowledge, which needs expertise of the environment, though. View-based navigation is helpful for (routes in) newly learned environments, for distances in which path integration and spatial updating would lead to poor results because the accumulated error would be too large, and for creating and establishing survey knowledge if several (crossing) routes have been learned. Again, it is important to note that these views are probably richer in features and are not simply composed of "mental photographs" as experiment 5 has shown.

A central structure in the brain that allows for such formidable navigation behavior is the hippocampus. Byrne et al. (2007) suggested that allocentric (long-term memory) representations were formed in the hippocampus, while spatial working memory is located in the precuneus, and that spatial information is transferred between these two. Courtney et al. (1998) identified an area in the superior frontal sulcus specialized for spatial working memory. A study by van Asselen et al. (2006) reported impaired performance in spatial memory tasks upon damage to the right posterior parietal cortex and right dorsolateral prefrontal cortex as well as to the hippocampus. Also in mental transformation of positions in a grid, the parietal cortex has been identified as a region of increased event related potentials in the brain (Rolke et al., 2000). Spiers & Maguire (2007) identified the posterior parietal cortex to be associated with route navigation as a correlation between its bilateral activation and the egocentric direction to goals confirmed. Furthermore, they found a correlation of goal proximity and activity of the medial prefrontal cortex. This structure is therefore supposed to interact with regionalization and account for the different response times in experiment 5 of this thesis as well as the lack of effect for two interview locations in experiment 4 a).
For navigation with landmarks, Epstein & Vass (2014) attributed a critical role to the parahippocampal place area. It is responsible for landmark recognition and representing "places by encoding the geometry of the local environment" (Epstein & Kanwisher, 1998). This area is supposed to fulfill a role that is analogous to the specialized face recognition areas, and it causes strong signals in fMRI upon presentation of environmental stimuli like streets, rooms, landscapes, etc. - even when they were viewed passively. Janzen & Turennout (2004) found that the activation of the parahippocampal place area also depends on the navigational relevance of objects. Interestingly, this area showed no response at all for faces, and only a weak response for everyday objects (Epstein & Vass, 2014).
The retrosplenial/medial parietal cortex is thought to be involved in determining one's current location and heading in the environment. It also shows increased activation in passive viewing of a scene, comparable to the parahippocampal place area; however, it strongly responds to scenes showing familiar locations (Epstein & Vass, 2014). Also Wolbers & Buchel (2005) reported a view specificity of the retrosplenial cortex. Patients with damage to this retrosplenial complex suffer from orientation problems in large-scale environments since they cannot use landmark information to orient themselves at a location (Aguirre & D'Esposito, 1999).
Furthermore, the medial temporal lobe, including the hippocampus, is supposed to encode an allocentric cognitive map with landmarks and goals, as Epstein & Vass (2015) but also other studies have found. Woollett & Maguire (2011), for instance, found an increased size of the hippocampi in taxi drivers, who have a magnificent knowledge of their city, and Schinazi et al. (2013) reported a correlation between the size of the right posterior hippocampus and the ability to memorize spatial layouts.
Also in other species, like rats, the hippocampus has been identified to play an important role in spatial memory (Clark et at., 2005), e.g. in solving Morris water maze tasks in which a location in a water tank has to be memorized and found by means of landmarks (Morris, 1984). It was also in

rats where so called “place cells” (O’Keefe & Dostrovsky, 1971) and “grid cells” (Hafting et al., 2005) were found first. Individual place cells fire the strongest if the rat is located at a certain spot in the (test) environment, and they are therefore thought to “code that place”. Grid cells exhibit a recurring increase and decrease in activation in a grid-like pattern while the rat is moving and could be used to track the distance traveled. See Moser et al. (2015) for a review.
The transition from route knowledge to the formation of a cognitive map was investigated by Brown et al. (2010). They found that in the hippocampus there were specific neurons that responded to overlapping navigational trajectories. This provides further evidence for the integration of individual (and overlapping) routes into a cognitive map. Finally, Lambrey et al. (2003) suggested a functional lateralization of the medial temporal lobe in which the left medial temporal lobe is supposed to be involved in sequential or route memory, while the right medial temporal lobe focuses on memory that holds “holistic” or map knowledge.
Altogether, in several studies different brain areas have been identified that are involved in spatial cognition and navigation, each of them fulfilling different roles. Two of these roles are associated with survey knowledge and route knowledge, which is further evidence that both representations can exist in parallel and survey knowledge does not rule out view-based navigation.

The experiments in this thesis were designed to investigate aspects of face and place recognition and representation. While being different at a first glance, they have several commonalities.
First, they both rely on interacting brain areas, which fulfill different roles, like basic recognition of features in faces and of (potential) landmarks, combining these features into holistic representations of a face and place, performing actions on/with these representations like identification of familiar ones, updating of changes (new looks; new position), and finally embedding in a network (semantic information for identities and relationships; view-graph and map). These brain areas are mostly found in the temporal lobe, however, in different areas within this structure. A study by Haxby et al. (2001) was investigating activation patterns in the ventral temporal cortex and they concluded that representations of faces and man-made objects (including the fronts of houses in their experiment) were widely distributed and overlapping. Additionally and depending on the task, other brain areas are involved, too, like the frontal lobe for (conscious) reasoning and working memory, the parietal lobe for egocentric spatial integration, or the occipital lobe for receiving (processed) visual information in general.
Second, both representations are prone to priming. In the first experiments 1 to 3, identity priming could be accomplished with pictures of faces and persons’ identities via names, however, with different magnitudes. In the latter experiments 4 and 5, spatial priming was achieved by mental imagery and also by showing pictures.
Third, there are special entities in both representations that are handled differently from the rest - those involve the “self”. For faces, a picture of oneself - and along with it the identity - has a special meaning. This uniqueness manifests in experimental results as could be seen in experiments 1 and 3 of this thesis, but also in different brain activation patterns in fMRI scans of other studies (see above) upon contrasting with several other stimuli. In spatial cognition, a similar effect can be found. The spot where someone is currently located is treated differently than other locations or when this person is somewhere else. Additionally, also the surrounding area is influenced by the presence as the local spatial networks (be it view-graph or map) get activated. This could be shown in both experiments 4 and 5.

Conclusion

Five experiments were accomplished in this thesis, investigating face and place representation. Faces are likely to be represented in a multidimensional face-space; however, they are not represented equally but with varying granularity. One's own face, or "self", forms a special instance here. This granularity leads to non-linear metrics within the face-space. Shifts occur naturally and are connected to ambiguity between faces. Shifts can also be induced by priming, which leads to changes in the perceived identity of a face. Priming effects with different stimuli are dependent on the task that the participants have to solve and do not exhibit a general pattern independent of the instructions.

The representation of places can also be primed by external and internal cues and is also dependent on one's current location. Analogous to a "fine-to-coarse" route planning strategy (Wiener et al., 2003), places in proximity to a navigator are reported ego-centrically, that is, they are oriented towards the navigator in the direction of approach, while distant places are reported in an allocentric way with a generic orientation. Furthermore, images of places in proximity to the navigator are recognized faster if they were taken in the direction of approach. Between the near scale "spatial image" and far scale "cognitive map", the view graph is an intermediate representation of the environment, especially for route knowledge.

The experiments of this thesis have shown that the "self" influences the shaping of these representations (faces and places) and influences the processing of face recognition and navigation.

11. References

Aguirre, G. K., & D'Esposito M. (1999). Topographical disorientation: a synthesis and taxonomy. *Brain*, 122(9), 1613-1628.

Allen, M., & Schwartz, B. L. (2008). Mirror Self-Recognition in a Gorilla (*Gorilla gorilla gorilla*). *Electronic Journal of Integrative Biosciences*, 5(1), 19-24.

Amsterdam, B. (1972). Mirror self-image reactions before age two. *Developmental Psychobiology*, 5(4), 297-305.

Anaki, D., Kaufman, Y., Freedman, M., & Moscovitch, M. (2007). Associative (prosop)agnosia without (apparent) perceptual deficits: a case-study. *Neuropsychologia*, 45 (8), 1658-1671. doi:10.1016/j.neuropsychologia.2007.01.003

Andreano, J. M., & Cahill, L. (2009). Sex influences on the neurobiology of learning and memory. *Learning & Memory*, 16, 248-266. doi: 10.1101/lm.918309

Armann, R., & Bülthoff, I. (2012). Male and female faces are only perceived categorically when linked to familiar identities - And when in doubt, he is a male. *Vision Research*, 63, 69-80.

Astur, R. S., Tropp, J., Sava, S., Constable, R. T., & Markus, E. J. (2004). Sex differences and correlations in a virtual Morris water task, a virtual radial arm maze, and mental rotation. *Behavioural Brain Research*, 151, 103-115.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science*, 4(11), 417-423.

Baddeley, A., & Hitch, G.J. (1974). Working memory. *Psychology of Learning and Motivation*, 8, 47-89.

Bartlett, J. C., Hurrey, S., & Thorley, W. (1984). Typicality and familiarity of faces. *Memory & Cognition*, 12, 219-228.

Basten, K., Meilinger, T., & Mallot, H. A. (2012). Mental travel primes place orientation in spatial recall. *Lecture Notes in Artificial Intelligence*, 7463, 378-385.

Bastin, C., & Van der Linden, M. (2003). The Contribution of Recollection and Familiarity to Recognition Memory: A Study of the Effects of Test Format and Aging. *Neuropsychology*, 17(1), 14-24.

Baudouin, J.-Y., & Brochard, R. (2011). Gender-Based Prototype Formation in Face Recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(4), 888-898.

Baudouin, J.-Y., & Gallay, M. (2006). Is Face Distinctiveness Gender Based? *Journal of Experimental Psychology: Human Perception and Performance*, 32(4), 789-798.

Batschelet, E. (1981). Circular Statistics in Biology. London: Academic Press.

Bayley, P. J., Wixted, J. T., Hopkins, R. O., & Squire, L. R. (2008). Yes/No Recognition, Forced-choice Recognition, and the Human Hippocampus. *Journal of Cognitive Neuroscience*, 20(3), 505-512.

Beale, J. M., & Keil, F. C. (1995). Categorical effects in the perception of faces. *Cognition*, 57, 217-239. doi: 10.1016/0010-0277(95)00669-X

11. References

Bekoff, M. (2001). Observations of scent-marking and discriminating self from others by a domestic dog (*Canis familiaris*): tales of displaced yellow snow. *Behavioural Processes*, 55(2), 75-79.

Berens, P. (2009). CircStat: A Matlab toolbox for circular statistics. *Journal of Statistical Software*, 31(10).

Bird, C. M., & Burgess, N. (2008). The hippocampus and memory: insights from spatial processing. *Nature Reviews Neuroscience*, 9(3), 182-194. doi: 10.1038/nrn2335

Birmelin, I. The self-recognition of an African elephant. *Verein für Verhaltensforschung bei Tieren e.V.*. http://www.tierverhaltensforschung-birmelin.de/aktuelles/the-self-recognition-of-an-african-elephant.html; [accessed 2019 Jan 31].

Bisiach, E., & Luzzatti, C. (1978). Unilateral neglect of representational space. *Cortex*, 14(1), 129-133.

Blackford, J. U., Allen, A. H., Cowan, R. L., & Avery, S. N. (2012). Amygdala and hippocampus fail to habituate to faces in individuals with an inhibited temperament. *Social Cognitive and Affective Neuroscience*, 8(2), 143-150.

Bodamer, J. (1947). Die Prosop-Agnosie. *Archiv für Psychiatrie und Nervenkrankheiten*, 179 (1-2), 6-53. doi: 10.1007/BF00352849

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, 113(4), 700-765.

Bornstein, M. H., & Korda, N. O. (1984). Discrimination and matching within and between hues measured by reaction times: Some implications for categorical perception and levels of information processing. *Psychological Research*, 47, 5-28.

Brown, T. I., Ross, R. S., Keller, J. B., Hasselmo, M. E., & Stern, C. E. (2010). Which Way Was I Going? Contextual Retrieval Supports the Disambiguation of Well Learned Overlapping Navigational Routes. *The Journal of Neuroscience*, 31(21), 7414-7422.

Bruce, V., Burton, A. M., & Dench, N. (1994). What's Distinctive About a Distinctive Face? *The Quarterly Journal of Experimental Psychology*, 47A (1), 119-141.

Bruce, V., & Valentine, T. (1986). Semantic priming of familiar faces. *The Quarterly Journal of Experimental Psychology*, Section A 38, 125-150.

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, 77,305-327.

Bulthoff, H. H., Edelman, S. Y., & Tarr, M. J. (1995). How are 3-dimensional objects represented in the brain. *Cerebral Cortex*, 5(3), 247-260. doi: 10.1093/cercor/5.3.247

Bülthoff, I. & Newell, F. N. (2004). Categorical perception of sex occurs in familiar but not unfamiliar faces. *Visual Cognition*, 11 (7), 823-855.

Bülthoff, I. & Newell, F. N. (2005). Categorical perception of gender: No evidence for unfamiliar faces. Max Planck Institute for Biological Cybernetics, Technical Report No. 094, http://www.kyb.tuebingen.mpg.de/fileadmin/user_upload/files/publications/techreport_094_%5B0%5D.pdf

Burgess, N. (2006). Spatial memory: how egocentric and allocentric combine. *Trends in Cognitive Sciences*, 10(12), 551-557. doi: 10.1016/j.tics.2006.10.005

Burton, A. M., & Bruce, V. (1993). Naming faces and naming names: Exploring an interactive activation model of person recognition. *Memory*, 1 (4), 457-480.

Burton, A. M., Bruce, V., & Hancock, P. J. B. (1999). From Pixels to People: A Model of Familiar Face Recognition. *Cognitive Science*, 23 (1), 1-31.

Burton, A. M., Bruce, V., & Johnston, R. A. (1990). Understanding face recognition with an interactive activation model. *British Journal of Psychology*, 81, 361-380.

Burton, A. M., & Vokey, J. R. (1998). The Face-Space Typicality Paradox: Understanding the Face-Space Metaphor. *The Quarterly Journal of Experimental Psychology*, 51A (3), 475-483. doi: 10.1080/713755768

Byrne, P., Becker, S., & Burgess, N. (2007). Remembering the past and imagining the future: A neural model of spatial memory and imagery. *Psychological Review*, 114(2), 340-375. doi: 10.1037/0033-295x.114.2.340

Calder, A. J., & Young, A. W. (1996). Self Priming: A Short-term Benefit of Repetition. *The Quarterly Journal of Experimental Psychology*, Section A 49, 845-861.

Calder, A. J., Young, A. W., Benson, P. J., & Perrett, D. I. (1996 a). Self priming from distinctive and caricatured faces. *British Journal of Psychology*, 87, 141-162.

Calder, A. J., Young, A. W., Perrett, D. I., Etcoff, N. L., & Rowland, D. (1996 b). Categorical perception of morphed facial expressions. *Visual Cognition*, 3 (2), 81-117.

Cammaerts, M.-C., & Cammaerts, R. (2015). Are ants (Hymenoptera, Formicidae) capable of self recognition? *Journal of Science*, 5(7), 521-532.

Campanella, S., Chrysochoos, A., & Bruyer, R. (2001). Categorical perception of facial gender information: Behavioural evidence and the face-space metaphor. *Visual Cognition*, 8 (2), 237-262.

Campanella, S., Hanoteau, C., Seron, X., Joassin, F., & Bruyer, R. (2003). Categorical perception of unfamiliar facial identities, the face-space metaphor, and the morphing technique. *Visual Cognition*, 10 (2), 129-156.

Carruthers, P. (2013). Evolution of working memory. *Proceedings of the National Academy of Sciences of the United States of America*, 110, 10371-10378. doi: 10.1073/pnas.1301195110

Cartwright, B. A., & Collett, T. S. (1982). How honey bees use landmarks to guide their return to a food source. *Nature*, 295(5850), 560-564. doi: 10.1038/295560a0

Chen, G. F., King, J. A., Burgess, N., & O'Keefe, J. (2013). How vision and movement combine in the hippocampal place code. *Proceedings of the National Academy of Sciences of the United States of America*, 110(1), 378-383. doi: 10.1073/pnas.1215834110

Chiroro, P., & Valentine, T. (1995). An investigation of the contact hypothesis of the own-race bias in face recognition. *The Quarterly Journal of Experimental Psychology*, 48(A), 879-894.

Clark, R. E., Broadbent, N. J., & Squire, L. R. (2005). Hippocampus and Remote Spatial Memory in Rats. *Hippocampus*, 15(2), 260-272.

Courtney S. M., Petit, L., Maisog, J. M., Ungerleider, L. G., & Haxby, J. V. (1998). An Area Specialized for Spatial Working Memory in Human Frontal Cortex. *Science*, 279(5355), 1347-1351.

Curby, K. M., Johnson, K. J., & Tyson, A. (2011). Face to face with emotion: Holistic face processing is modulated by emotional state. *Cognition and Emotion*, 26(1), 93-102.

Davies-Thompson, J., Newling, K., & Andrews, T. J. (2012). Image-Invariant Responses in Face-Selective Regions Do Not Explain the Perceptual Advantage for Familiar Face Recognition. *Cerebral Cortex*, 23(2), 370-377.

De Renzi, E., & di Pellegrino, G. (1998). Prosopagnosia and Alexia Without Object Agnosia. *Cortex*, 34 (3), 403-415. doi: 10.1016/S0010-9452(08)70763-9

Deffenbacher, K. A., Vetter, T., Johanson, J., & O'Toole, A. J. (1998). Facial Aging, Attractiveness, and Distinctiveness. *Perception*, 27 (10), 1233-1243. doi: 10.1068/p271233

Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J.-F., Poline, J.-B., & Rivière, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience*, 4(7), 752-758.

Delfour, F., & Marten, K. (2001). Mirror image processing in three marine mammal species: killer whales (*Orcinus orca*), false killer whales (*Pseudorca crassidens*) and California sea lions (*Zalophus californianus*). *Behavioural Processes*, 53(3), 181-190.

Dell'Acqua, R., & Grainger, J. (1999). Unconscious semantic priming from pictures. *Cognition* 73, B1-B15.

Denis, M. (1997). The description of routes: A cognitive approach to the production of spatial discourse. *Current Psychology of Cognition*, 16 (4), 409-458.

Devue, C., & Brédart, S. (2011). The neural correlates of visual self-recognition. *Consciousness and Cognition*, 20, 40-51.

Eimer, M., & Schlaghecken, F. (2003). Response facilitation and inhibition in subliminal priming. *Biological Psychology*, 64, 7-26.

Ellis, A. W., Young, A. W., Flude, B. M., & Hay, D. (1987). Repetition priming of face recognition. *The Quarterly Journal of Experimental Psychology*, Section A 39, 193-210.

Enlow, D. H. (1990). *Facial growth* (3rd ed.). Philadelphia, PA: Harcourt Brace Jovanovich.

Epstein, R. A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends in Cognitive Sciences*, 12(10), 388-396. doi: 10.1016/j.tics.2008.07.004

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, 392(6676), 598-601.

Epstein, R. A., & Vass, L. K. (2014). Neural systems for landmark-based wayfinding in humans. Philosophical Transactions of the Royal Society B, 369: 20120533.

Etcoff, N. L., & Magee, J. J. (1992). Categorical perception of facial expressions. *Cognition*, 44 (3), 227-240. doi: 10.1016/0010-0277(92)90002-Y

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review*, 105 (3), 482-498. doi: 10.1037/0033-295X.105.3.482

Farrell, M. J., & Robertson, I. H. (1998). Mental rotation and the automatic updating of body-centered spatial relationships. *Journal of Experimental Psychology-Learning Memory and Cognition*, 24(1), 227-233. doi: 10.1037//0278-7393.24.1.227

Field, T. M., Woodson, R., Greenberg, R., & Cohen, D. (1982). Discrimination and imitation of facial expression by neonates. *Science*, 218(4568), 179-181.

Forgas, J. P. (1995). Mood and judgment: The affect infusion model (AIM). *Psychological Bulletin*, 117(1), 39-66.

Franz, M. O., Scholkopf, B., Mallot, H. A., & Bulthoff, H. H. (1998). Learning view graphs for robot navigation. *Autonomous Robots*, 5(1), 111-125. doi: 10.1023/a:1008821210922

Freeman, J. B., Stolier, R. M., Ingbretsen, Z. A., & Hehman, E. A. (2014). Amygdala Responsivity to High-Level Social Information from Unseen Faces. *The Journal of Neuroscience*, 34(32), 10573-10581.

Fried, I., MacDonald, K. A., & Wilson, C. L. (1997). Single Neuron Activity in Human Hippocampus and Amygdala during Recognition of Faces and Objects. *Neuron*, 18, 753-765.

Furuya, Y., Matsumoto, J., Hori, E., Boas, C. V., Tran, A. H., Shimada, Y., Ono, T.,& Nishijo, H. (2014). Place-Related Neuronal Activity in the Monkey Parahippocampal Gyrus and Hippocampal Formation During Virtual Navigation. *Hippocampus*, 24(1), 113-130. doi: 10.1002/hipo.22209

Galea, L. A. M., & Kimura, D. (1993). Sex Differences in Route-Learning. *Personality and Individual Differences*, 14(1), 53-65. doi: 10.1016/0191-8869(93)90174-2

Gallup Jr., G. G. (1970). Chimpanzees: Self-Recognition. *Science*, 167(3914), 86-87.

Ganel, T., & Goshen-Gottstein, Y. (2002). Perceptual integrality of sex and identity of faces: further evidence for the single-route hypothesis. *Journal of Experimental Psychology: Human Perception and Performance*, 28(4), 854-867.

Gaussier, P., Revel, A., Banquet, J. P., & Babeau, V. (2002). From view cells and place cells to cognitive map learning: processing stages of the hippocampal system. *Biological Cybernetics*, 86(1), 15-28. doi: 10.1007/s004220100269

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, 3, 191-197.

Ghuman, A. S., Brunet, N. M., Li, Y., Konecky, R. O., Pyles, J. A., Walls, S. A., Destefino, V., Wang, W., & Richardson, R. M. (2014). Dynamic encoding of face information in the human fusiform gyrus. *Nature Communications*, 5, article number 5671. doi: 10.1038/ncomms6672

Gillner, S., Weiss, A. M., & Mallot, H. A. (2008). Visual homing in the absence of feature-based landmark information. *Cognition*, 109(1), 105-122. doi: 10.1016/j.cognition.2008.07.018

Giudice, N. A., Klatzky, R. L., Bennett, C. R., & Loomis, J. M. (2013). Combining Locations from Working Memory and Long-Term Memory into a Common Spatial Image. *Spatial Cognition and Computation*, 13(2), 103-128. doi: 10.1080/13875868.2012.678522

Golledge, R. G., Dougherty, V., & Bell, S. (1995). Acquiring Spatial Knowledge: Survey Versus Route-Based Knowledge in Unfamiliar Environments. *Annals of the Association of American Geographers*, 85(1), 134-158.

Goshen-Gottstein, Y., & Ganel, T. (2000). Repetition priming for familiar and unfamiliar faces in a sex-judgment task: evidence for a common route for the processing of sex and identity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(5), 1198-1214.

Graf, P., & Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(3), 501-518.

Gross, T. F. (2009). Own-Ethnicity Bias in the Recognition of Black, East Asian, Hispanic, and White Faces. *Basic and Applied Social Psychology*, 31 (2), 128-135.

Grusser, O. J. (1982). Space perception and the gazemotor system. *Human neurobiology*, 1(2), 73-76.

Hadjikhani, N., Kveraga, K., Naik, P., & Ahlfors, S. P. (2009). Early (M170) activation of face-specific cortex by face-like objects. *NeuroReport*, 20(4), 403-407.

Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., & Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature*, 436, 801-806.

Hardiess, G., Basten, K., & Mallot, H. A. (2011). Acquisition vs. Memorization Trade-Offs Are Modulated by Walking Distance and Pattern Complexity in a Large-Scale Copying Paradigm. *PLOS ONE*, 6(4). doi: 10.1371/journal.pone.0018494

Harnad, S. (1987). Chapter 1: Psychological and cognitive aspects of categorical perception: A critical overview. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition*. New York: Cambridge University Press.

Hartley, T., Maguire, E. A., Spiers, H. J., & Burgess, N. (2003). The well-worn route and the path less traveled: Distinct neural bases of route following and wayfinding in humans. *Neuron*, 37(5), 877-888. doi: 10.1016/s0896-6273(03)00095-3

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and Overlapping Representations of Faces and Objects in Ventral Temporal Cortex. *Science*, 293(5539), 2425-2430.

Haxby, J. V., Hoffmann, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, 4(6), 223-233.

Heil, M., Kavsek, M., Rolke, B., Beste, C., & Jansen, P. (2011). Mental rotation in female fraternal twins: Evidence for intra-uterine hormone transfer? *Biological Psychology*, 86(1), 90-93.

Heit, G., Smith, M. E., & Halgren, E. (1988). Neural encoding of individual words and faces by the human hippocampus and amygdala. *Nature*, 333, 773-775.

Hoehl, S., & Striano, T. (2008). Neural Processing of Eye Gaze and Threat-Related Emotional Facial Expressions in Infancy. *Child Development*, 79(6), 1752-1760.

Hölscher, C., Tenbrink, T., & Wiener, J. (2011). Would you follow your own route description? Cognitive strategies in urban route planning. *Cognition*, 121, 228-247.

Hooven, C. K., Chabris, C. F., Ellison, P. T., & Kosslyn, S. M. (2004). The relationship of male testosterone to components of mental rotation. *Neuropsychologia*, 42(6), 782-790.

Horowitz, A. (2017). Smelling themselves: Dogs investigate their own odours longer when modified in an “olfactory mirror” test. *Behavioural Processes*, 143, 17-24.

Ison, M. J., Mormann, F., Cerf, M., Koch, C., Fried, I., & Quiroga, R. Q. (2011). Selectivity of pyramidal cells and interneurons in the human medial temporal lobe. *Journal of Neurophysiology*, 106(4), 1713-1721. doi: 10.1152/jn.00576.2010

Jacobs, W. J., Thomas, K. G. F., Laurance, H. E., & Nadel, L. (1998). Place Learning in Virtual Space: II. Topographical Relations as One Dimension of Stimulus Control. *Learning and Motivation*, 29(3), 288-308.

Janzen, G., & van Turennout, M. (2004). Selective neural representation of objects relevant for navigation. *Nature Neuroscience*, 7, 673-677.

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society B*, 361, 2109-2128.

Keenan, J. P., Wheeler, M. A., Gallup Jr., G. G., & Pascual-Leone, A. (2000). Self-recognition and the right prefrontal cortex. *Trends in Cognitive Sciences*, 4(9), 338-344.

Kircher, T. T. J., Senior, C., Phillips, M. L., Rabe-Hesketh, S., Benson, P. J., Bullmore, E. T., Brammer, M., Simmons, A., Bartels, M., & David, A. S. (2001). Recognizing one's own face. *Cognition*, 78, B1-B15.

Klatzky, R. L. (1998). Allocentric and egocentric spatial representations: Definitions, distinctions and interconnections. *Lecture Notes in Artificial Intelligence*, 1404, 2 - 17.

Klatzky, R. L., & Wu, B. (2008). The Embodied Actor in Multiple Frames of Reference. In R. L. Klatzky, B. MacWhinney, & M. Behrman (Eds.), *Carnegie Mellon symposia on cognition. Embodiment, ego-space, and action* (pp. 145-177). New York: Psychology Press.

Kozlowski, L. T., & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Attention, Perception, & Psychophysics*, 21(6), 575-580.

Lambrey, S., Doeller, C., Berthoz, A., & Burgess, N. (2012). Imagining Being Somewhere Else: Neural Basis of Changing Perspective in Space. *Cerebral Cortex*, 22(1), 166-174. doi: 10.1093/cercor/bhr101

Lambrey, S., Samson, S., Dupont, S., Baulac, M., & Berthoz, A. (2003). Reference frames and cognitive strategies during navigation: is the left hippocampal formation involved in the sequential aspects of route memory? *International Congress Series*, 1250, 261-274. doi:10.1016/S0531-5131(03)00997-X

Lancier, S. (2016). Spatial memories in place recognition. *Doctoral thesis*. Department of Biology, Intitute for Neurobiology, Eberhard Karls Universität Tübingen.

Land, M. F. (2014). Do we have an internal model of the outside world? *Philosophical Transactions of the Royal Society B*, 369(1636). doi: 10.1098/rstb.2013.0045

Leppänen, J. M., Moulson, M. C., Vogel-Farley, V. K., & Nelson, C. A. (2007). An ERP Study of Emotional Face Processing in the Adult and Infant Brain. *Child Development*, 78(1), 232-245.

Leube, D. T., Erb, M., Grodd, W., Bartels, M., & Kircher, T. T. J. (2003). Successful episodic retrieval of newly learned faces activates a left fronto-parietal network. *Cognitive Brain Research*, 18, 97-101.

Leveroni, C. J., Seidenberg, M., Mayer, A. R., Mead, L. A., Binder, J. R., & Rao, S. M. (2000). Neural Systems Underlying the Recognition of Familiar and Newly Learned Faces. *The Journal of Neuroscience*, 20(2), 878-886.

Levin, D. T., & Angelone, B. L. (2002). Categorical perception of race. *Perception*, 31 (5), 567-578. doi: 10.1068/p3315

Levin, D. T., & Beale, J. M. (2000). Categorical perception occurs in newly learned faces, other-race faces, and inverted faces. *Perception & Psychophysics*, 62 (2), 386-401.

Lewis, M. B., & Johnston, R. A. (1999). A Unified Account of the Effects of Caricaturing Faces. *Visual Cognition*, 6(1), 1-41.

Lieberman, H. R., & Pentland, A. P. (1982). Microcomputer-based estimation of psychophysical thresholds: The Best PEST. *Behavior Research Methods & Instrumentation*, 14 (1), 21-25.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74, 431-461.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 54, 358-368.

Light, L. L., Kayra-Stuart, F., & Hollander, S. (1979). Recognition memory for typical and unusual faces. *Journal of Experimental Psychology: Human Learning and Memory*, 5, 212-228.

Little, A. C., & Hancock, P. J. B. (2002). The role of masculinity and distinctiveness in judgments of human male facial attractiveness. *British Journal of Psychology*, 93, 451-464.

Liu, J., Harris, A., & Kanwisher, N. (2010). Perception of Face Parts and Face Configurations: An fMRI Study. *Journal of Cognitive Neuroscience*, 22(1), 203-211.

Lohmann, J., Rolke, B., & Butz, M. V. (2017). In touch with mental rotation: interactions between mental and tactile rotations and motor responses. *Experimental Brain Research*, 234(4), 1063-1079.

Loomis, J. M., Klatzky, R. L., & Giudice, N. A. (2013). Representing 3D space in working memory: Spatial images from vision, hearing, touch, and language. In S. Lacey & R. Lawson (Eds.), *Multisensory Imagery*. New York: Springer.

Loomis, J. M., Klatzky, R. L., Golledge, R. G., Cicinelli, J. G., Pellegrino, J. W., & Fry, P. A. (1993). Nonvisual navigation by blind and sighted - Assessment of path integration ability. *Journal of Experimental Psychology: General*, 122(1), 73-91. doi: 10.1037/0096-3445.122.1.73

Lovelace, K. L., Hegarty, M., & Montello, D. R. (1999). Elements of Good Route Directions on Familiar and Unfamiliar Environments. In C. Freksa & D. M. Mark (Eds.), *Spatial Information Theory*, 1661, 65-82. Berlin: Springer.

Lucas, M. (2000). Semantic priming without association. A meta-analytic review. *Psychonomic Bulletin and Review*, 7(4), 618-630.

Malinowski, J. C., & Gillespie, W. T. (2001). Individual differences in performance on a large-scale real-world wayfinding task. *Journal of Environmental Psychology*, 21, 73-82.

Mallot, H. A. (2000). Sehen und die Verarbeitung visueller Information. 2. Aufl., Kap. 11, Braunschweig/Wiesbaden: Fried. Vieweg & Sohn Verlagsgesellschaft mbH.

Mallot, H. A., & Basten, K. (2009). Embodied spatial cognition: Biological and artificial systems. *Image and Vision Computing*, 27(11), 1658-1670. doi: 10.1016/j.imavis.2008.09.001

Mallot, H. A., & Gillner, S. (2000). Route navigating without place recognition: What is recognised in recognition-triggered responses? *Perception*, 29(1), 43-55. doi: 10.1068/p2865

Marten, K. & Psarakos, S. (1994). Evidence of self-awareness in the bottlenose dolphin (*Tursiops truncatus*). In S. T. Parker, R. W. Mitchell & M. L. Boccia (Eds.), *Self-awareness in animals and humans: Developmental perspectives*, 361-379. Cambridge University Press.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, 88(5), 375-407.

McGee, M. G. (1979). Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences. *Psychological Bulletin*, 86, 889-918.

McGuiness, D., & Sparks, J. (1983). Cognitive style and cognitive maps: Sex differences in representations of a familiar terrain. *Journal of Mental Imagery*, 7, 91-100.

Meilinger, T., Frankenstein, J., Simon, N., Bülthoff, H. H., & Bresciani, J.-P. (2016). Not all memories are the same: Situational context influences spatial recall within one’s city of residency. *Psychonomic Bulletin and Review*, 23(1), 246-252.

Meilinger, T., Frankenstein, J., Watanabe, K., Bülthoff, H. H., & Hölscher, C. (2015). Reference frames in learning from maps and navigation. *Psychological Research*, 76(6), 1000-1008. doi: 10.1007/s00426-014-0629-6

Meilinger, T., Knauff, M., & Bulthoff, H. H. (2008). Working memory in wayfinding - A dual task experiment in a virtual city. *Cognitive Science*, 32(4), 755-770. doi: 10.1080/03640210802067004

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law*, 7(1), 3-35.

Mellet, E., Bricogne, S., Crivello, F., Mazoyer, B., Denis, M., & Tzourio-Mazoyer, N. (2002). Neural basis of mental scanning of a topographic representation built from a text. *Cerebral Cortex*, 12(12), 1322-1330. doi: 10.1093/cercor/12.12.1322

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 90, 227-234.

Miller, L. K., & Santoni, V. (1986). Sex differences in spatial abilities: Strategic and experiential correlates. *Acta Psychologica*, 62, 225-235.

Moffat, S. D., Hampson, E., & Hatzipantelis, M. (1998). Navigation in a “Virtual” Maze: Sex Differences and Correlation With Psychometric Measures of Spatial Ability in Humans. *Evolution and Human Behavior*, 19, 73-87.

Moller, R., Krzykawski, M., & Gerstmayr, L. (2010). Three 2D-warping schemes for visual robot navigation. *Autonomous Robots*, 29(3-4), 253-291. doi: 10.1007/s10514-010-9195-y

Montello, D. R. (1993). Scale and multiple psychologies of space. *Lecture Notes in Computer Science*, 716, 312-321

Morris, R. (1984). Developments of a water-maze procedure for studying spatial learning in the rat. *Journal of Neuroscience Methods*, 11(1), 47-60.

Moser, M.-B., Rowland, D. C., & Moser, E. I. (2015). Place Cells, Grid Cells, and Memory. *Perspectives in Biology*, Cold Spring Harb Perspect Biol 2015;7:a021808. doi: 10.1101/cshperspect.a021808

Mou, W. M., & McNamara, T. P. (2002). Intrinsic frames of reference in spatial memory. *Journal of Experimental Psychology-Learning Memory and Cognition*, 28(1), 162-170. doi: 10.1037//0278-7393.28.1.162

Murphy, S. T., & Zajonc, R. B. (1993). Affect, Cognition, and Awareness: Affective Priming With Optimal and Suboptimal Stimulus Exposures. *Journal of Personality and Social Psychology*, 64(5), 723-739.

Nasr, S., Devaney, K. J., & Tootell, R. B. H. (2013). Spatial encoding and underlying circuitry in scene-selective cortex. *Neuroimage*, 83, 892-900. doi: 10.1016/j.neuroimage.2013.07.030

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition*, 264-336. Hillsdale, NJ: Erlbaum.

Nestor, A., Plaut, D. C., & Behrmann, M. (2011). Unraveling the distributed neural code of facial identity through spatiotemporal pattern analysis. *Proceedings of the National Academy of Sciences of the United States of America*, 108(24), 9998-10003.

O’Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, 34(1), 171-175.

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. London, UK: Oxford University Press.

Pannese, A., & Hirsch, J. (2010). Self-specific priming effect. *Consciousness and Cognition*, 19, 962-968.

Pearce, P. L. (1981). Route maps: A study of travellers' perceptions of a section of countryside. *Journal of Environmental Psychology*, 1(2), 141-155.

Pentland, A. (1980). Maximum likelihood estimation: The best PEST. *Perception & Psychophysics*, 28 (4), 377-379.

Philbeck, J. W., & Loomis, J. M. (1997). Comparison of two indicators of perceived egocentric distance under full-cue and reduced-cue conditions. *Journal of Experimental Psychology-Human Perception and Performance*, 23(1), 72-85. doi: 10.1037/0096-1523.23.1.72

Piccardi, L., Bianchini, F., Argento, O., De Nigris, A., Maialetti, A., Palermo, L., & Guariglia, C. (2013). The Walking Corsi Test (WalCT): standardization of the topographical memory test in an Italian population. *Neurological Sciences*, 34(6), 971-978. doi: 10.1007/s10072-012-1175-x

Pickup, L. C., Fitzgibbon, A. W., & Glennerster, A. (2013). Modelling human visual navigation using multi-view scene reconstruction. *Biological Cybernetics*, 107(4), 449-464. doi: 10.1007/s00422-013-0558-2

Pitcher, D., Walsh, V, & Duchaine, B. (2011). The role of the occipital face area in the cortical face perception network. *Experimental Brain Research*, 209(4), 481-493.

Platek, S. M., Thomson, J. W., & Gallup Jr, G. G. (2004). Cross-modal self-recognition: The role of visual, auditory, and olfactory primes. *Consciousness and Cognition*, 13(1), 197-210.

Plotnik, J. M., de Waal, F. B. M., & Reiss, D. (2006). Self-recognition in an Asian elephant. *Proceedings of the National Academy of Sciences of the United States of America*, 103(45), 17053-17057.

Posada, S., & Colell, M. (2006). Another gorilla (*Gorilla gorilla gorilla*) recognizes himself in a mirror. *American Journal of Primatology*, 69(5), 576-583.

Prior, H., Schwarz, A., & Güntürkün, O. (2008). Mirror-Induced Behavior in the Magpie (*Pica pica*): Evidence of Self-Recognition. *PLOS Biology*, 6(8), 1642-1650.

Ranganath, C., & Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. *Nature Reviews Neuroscience*, 13(10), 713-726. doi: 10.1038/nrn3338

Restat, J. D., Steck, S. D., Mochnatzki, H. F., & Mallot, H. A. (2004). Geographical slant facilitates navigation and orientation in virtual environments. *Perception*, 33, 667-687. doi:10.1068/p5030

Rhodes, G., Hickford, C., & Jeffery, L. (2000). Sex-typicality and attractiveness: Are supermale and superfemale faces super-attractive? *British Journal of Psychology*, 91(1), 125-140. doi: 10.1348/000712600161718

Roedinger, H. L. 3rd (1990). Implicit memory. Retention without remembering. *American Psychologist*, 45(9), 1043-1056.

Röhrich, W. G., Hardiess, G., & Mallot, H. A. (2014). View-Based Organization and Interplay of Spatial Working and Long-Term Memories. *PLOS ONE*, 9(11), e112793. doi: 10.1371/journal.pone.0112793

Rolke, B., Heil, M., Hennighausen, E., Häussler, C., & Rösler, F., (2000). Topography of brain electrical activity dissociates the sequential order transformation of verbal versus spatial information in humans. *Neuroscience Letters*, 282, 81-84.

Röser, A., Hardiess, G., & Mallot, H.A. (2016). Modality dependence and intermodal transfer in the Corsi Spatial Sequence Task: Screen vs. Floor. *Experimental Brain Research*, 234(7), 1849-1862. doi: 10.1007/s00221-016-4582-z

Rossano, M. J., & Reardon, W. P. (1999). Goal Specificity and the Acquisition of Survey Knowledge. *Environment and Behavior*, 31(3), 395-412.

Rossano, M. J., West, S. O., Robertson, T. J., Wayne, M. C., & Chase, R. B. (1999). The acquisition of route and survey knowledge from computer models. *Journal of Environmental Psychology*, 19(2), 101-115.

Rossion, B., Hanseeuw, B., & Dricot, L. (2012). Defining face perception areas in the human brain: A large-scale factorial fMRI face localizer analysis. *Brain and Cognition*, 79, 138-157.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(3), 501-518.

Schaeffer, B., & Wallace, R. (1969). Semantic similarity and the comparison of word meanings. *Journal of Experimental Psychology*, 82, 343-346.

Schinazi, V. R., Nardi, D., Newcombe, N. S., Shipley, T. F., & Epstein, R. A. (2013). Hippocampal size predicts rapid learning of a cognitive map in humans. *Hippocampus*, 23(6), 515-528.

Schindler, A., & Bartels, A. (2013). Parietal Cortex Codes for Egocentric Space beyond the Field of View. *Current Biology*, 23(2), 177-182. doi: 10.1016/j.cub.2012.11.060

Schölkopf, B., & Mallot, H. A. (1995). View-based cognitive mapping and path planning. *Adaptive Behavior*, 3(3), 311-348. doi: 10.1177/105971239500300303

Spiers, H. J., & Maguire, E. A. (2007). A navigational guidance system in the human brain. *Hippocampus*, 17(8), 618-626.

Striano, T., & Vaish, A. (2006). Seven- to 9-month-old infants use facial expressions to interpret others' actions. *British Journal of Developmental Psychology*, 24(4), 753-760.

Suárez, S. D., & Gallup Jr., G. G. (1981). Self-recognition in chimpanzees and orangutans, but not gorillas. *Journal of Human Evolution*, 10(2), 175-188.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology*, A 46 (2), 225-245.

Tatler, B. W., & Land, M. F. (2011). Vision and the representation of the surroundings in spatial memory. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 366(1564), 596-610. doi: 10.1098/rstb.2010.0188

Taubert, J., Apthorp, D., Aagten-Murphy, D., & Alais, D. (2011). The role of holistic processing in face perception: Evidence from the face inversion effect. *Vision Research*, 51 (11), 1273-1278.

Taylor, M. J., Arsalidou, M., Bayless, S. J., Morris, D., Evans, J. W., & Barbeau, E. J. (2009). Neural Correlates of Personally Familiar Faces: Parents, Partner and Own Faces. *Human Brain Mapping*, 30, 2008-2020.

Thorndyke, P. W., & Hayes-Roth, B. (1982). Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology*, 14(4), 560-589.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, 55(4), 189-208.

Tong, F., & Nakayama, K. (1999). Robust Representations for Faces: Evidence From Visual Search. *Journal of Experimental Psychology and Performance*, 25 (4), 1016-1035.

Ullman, S., & Basri, R. (1991). Recognition by linear-combinations of models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(10), 992-1006. doi: 10.1109/34.99234

Usher, M., & McClelland, J. J. (2001). The Time Course of Perceptual Choice: The Leaky, Competing Accumulator Model. *Psychological Review*, 108 (3), 550-592.

Valentine, T. (1991). A unified account on the effects of distinctiveness, inversion, and race in face recognition. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 43(A), 161-204.

Valentine, T., & Bruce, V. (1986). The effect of distinctiveness in recognizing and classifying faces. *Perception*, 15, 525-535.

Valentine, T., & Endo, M. (1992). Towards an exemplar model of face processing: The effects of race and distinctiveness. *Quarterly Journal of Experimental Psychology*, 44(A), 671-703.

Valentine, T., Lewis, M. B., & Hills, P. J. (2016). Face-space: A unifying concept in face recognition research. *The Quarterly Journal of Experimental Psychology*, 69(10), 1996-2019. doi: 10.1080/17470218.2014.990392

van Asselen, M., Kessels, R. P., Neggers, S. F., Kappelle, L. J., Frijns, C. J., & Postma, A. (2006). Brain areas involved in spatial working memory. *Neuropsychologia*, 44(7), 1185-1194.

Vann, S. D., Aggleton, J. P., & Maguire, E. A. (2009). What does the retrosplenial cortex do? *Nature Reviews Neuroscience*, 10(11), 792-U750. doi: 10.1038/nrn2733

Vorberg, D., Mattler, U., Heinecke, A., Schmidt, T., & Schwarzbach, J. (2003). Different time courses for visual perception and action priming. *Proceedings of the National Academy of Sciences of the United States of America*, 100(10), 6275-6280. doi: 10.1073/pnas.0931489100

Walraven, V., van Elsacker, L., & Verheyen, R. (1995). Reactions of a group of pygmy chimpanzees (*Pan paniscus*) to their mirror-images: Evidence of self-recognition. *Primates*, 36(1), 145-150.

Wang, R. F., & Brockmole, J. R. (2003). Simultaneous spatial updating in nested environments. *Psychonomic Bulletin & Review*, 10(4), 981-986. doi: 10.3758/bf03196562

Wang, R. F., & Spelke, E. S. (2002). Human spatial representation: Insights from animals. *Trends in Cognitive Sciences*, 6(9), 376-382. doi: 10.1016/s1364-6613(02)01961-7

Wang, R. X. F., & Brockmole, J. R. (2003). Human navigation in nested environments. *Journal of Experimental Psychology-Learning Memory and Cognition*, 29(3), 398-404. doi: 10.1037/0278-7393.29.3.398

Wei, W., Chen, C., Dong, Q., & Zhou, X. (2016). Sex Differences in Gray Matter Volume of the Right Anterior Hippocampus Explain Sex Differences in Three-Dimensional Mental Rotation. *Frontiers in Human Neuroscience*, 10, 1-9. doi: 10.3389/fnhum.2016.00580

Wiener, J. M., & Mallot, H. A. (2003). 'Fine-to-Coarse' Route Planning and Navigation in Regionalized Environments. *Spatial Cognition and Computation*, 3(4), 331-358.

Wiener, J. M., Schnee, A., & Mallot, H. A. (2004). Use and interaction of navigation strategies in regionalized environments. *Journal of Environmental Psychology*, 24, 475-493.

Wolbers, T., & Buchel, C. (2005). Dissociable retrosplenial and hippocampal contributions to successful formation of survey representations. *Journal of Neuroscience*, 25(13), 3333-3340. doi: 10.1523/jneurosci.4705-04.2005

Wolbers, T., Wiener, J. M., Mallot, H. A., & Buchel, C. (2007). Differential recruitment of the hippocampus, medial prefrontal cortex, and the human motion complex during path integration in humans. *Journal of Neuroscience*, 27(35), 9408-9416. doi: 10.1523/jneurosci.2146-07.2007

Woollett, K., & Maguire, E. A. (2011). Acquiring "the Knowledge" of London's Layout Drives Structural Brain Changes. *Current Biology*, 21, 2109-2114.

Zajonc, R. B. (1980). Feeling and Thinking: Preferences Need No Inferences. *American Psychologist*, 35, 151-175.

Zhang, J., Li, X., Song, Y., & Liu, J. (2012). The fusiform face area is engaged in holistic, not parts-based, representation of faces. *PLOS ONE*, 7(7), e40390. doi:10.1371/journal.pone.0040390

11. References

12. Appendix

12. Appendix

Experiment 1 - Discrimination and categorization of faces

Table 2. ANOVA table of pairwise comparisons of pooled response time data. Displayed are the morph-levels and their difference to 50:50 morph-level in terms of means and standard error as well as the probability of error (p-value). Significant p-values were written in bold, and morph-levels whose p-value was lower than 0.05 were marked with a star.

morph-level	difference of means	standard error	p-value
100:0 *	0.701	0.242	**0.018**
90:10 *	0.731	0.247	**0.016**
80:20 *	0.735	0.248	**0.016**
70:30	0.499	0.249	0.076
60:40	0.236	0.144	0.135
50:50	0	0	1
40:60	0.137	0.083	0.133
30:70 *	0.508	0.197	**0.030**
20:80 *	0.633	0.219	**0.018**
10:90 *	0.723	0.239	**0.014**
0:100 *	0.712	0.249	**0.019**

Experiment 2 - Face-space

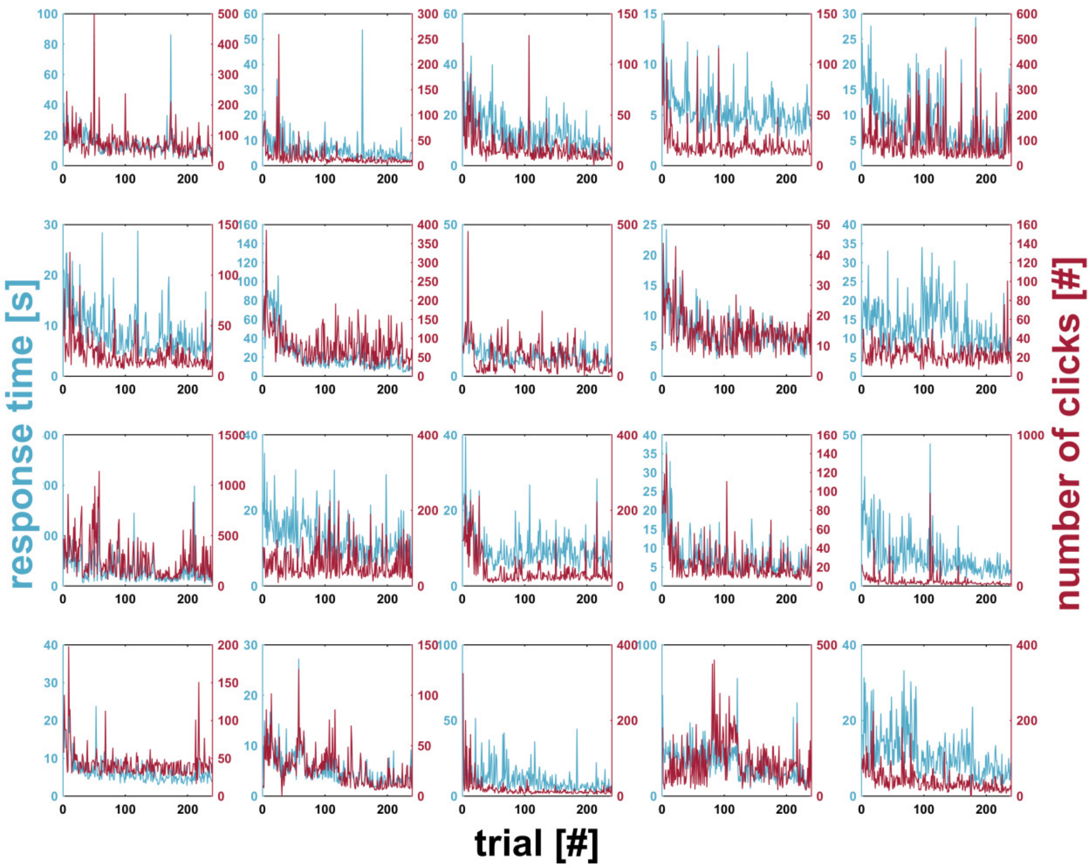

Figure 50. Individual response times and number of clicks across trials for all 20 participants. The x-axes show the trial, the y-axes show the response times in seconds (blue, left side) and number of clicks (red, right side). Please note the different scales.

Table 3 . Individual statistics per participant and face pair. Degrees of freedom were 11 for all t-tests, numbers in bold indicate significant p-values below 0.05. A mean of 13.5 equals to 50:50 % morph level. Smaller numbers indicate a shift towards the first face, larger numbers towards the second face.

VP	image pair	AB	AC	AD	AE	BC	BD	BE	CD	CE	DE	VW	VX	VY	VZ	WX	WY	WZ	XY	XZ	YZ
1	mean	14,25	14,08	14,17	13,42	13,08	14,58	13,50	13,00	12,50	13,58	13,33	12,75	14,17	13,50	13,58	14,33	12,25	14,83	12,92	14,33
	SD	1,42	1,51	1,85	1,38	1,38	1,88	2,75	2,76	1,51	1,88	1,61	1,36	2,08	1,24	2,35	2,06	1,36	2,08	1,56	1,67
	t-value	1,827	1,343	1,248	-0,209	-1,047	1,995	0,000	-0,627	-2,298	0,153	-0,358	-1,915	1,109	0,000	0,123	1,402	-3,191	2,219	-1,292	1,729
	p-value	0,095	0,206	0,238	0,838	0,318	0,071	1,000	0,544	**0,042**	0,881	0,727	0,082	0,291	1,000	0,905	0,189	**0,009**	**0,048**	0,223	0,112
2	mean	12,17	12,42	14,33	12,75	12,67	13,92	14,25	13,50	13,67	14,17	14,08	13,00	13,75	13,83	13,92	13,42	14,00	14,83	13,58	11,75
	sd	1,95	1,93	2,39	2,26	1,15	2,78	3,02	2,97	1,87	1,47	1,73	2,00	1,06	1,80	1,88	2,54	2,56	1,34	2,23	2,18
	t-value	-2,373	-1,946	1,209	-1,149	-2,500	0,519	0,861	0,000	0,308	1,574	1,168	-0,866	0,821	0,641	0,767	-0,114	0,677	3,454	0,129	-2,782
	p-value	**0,037**	0,078	0,252	0,275	**0,030**	0,614	0,408	1,000	0,764	0,144	0,267	0,405	0,429	0,534	0,459	0,912	0,512	**0,005**	0,900	**0,018**
3	mean	11,42	12,83	12,58	14,00	12,00	13,33	14,33	12,67	13,17	14,08	12,17	14,00	13,08	13,25	14,58	13,25	12,67	12,58	13,25	13,42
	sd	1,78	2,89	2,15	2,37	1,95	1,72	1,87	2,42	2,41	1,68	1,64	2,49	1,31	1,14	2,02	1,91	2,19	1,31	1,60	1,08
	t-value	-4,051	-0,800	-1,476	0,730	-2,659	-0,335	1,540	-1,191	-0,480	1,205	-2,812	0,697	-1,101	-0,761	1,857	-0,453	-1,319	-2,421	-0,540	-0,266
	p-value	**0,002**	0,441	0,168	0,481	**0,022**	0,744	0,152	0,259	0,641	0,253	**0,017**	0,500	0,295	0,463	0,090	0,660	0,214	**0,034**	0,600	0,795
4	mean	13,83	12,92	14,75	13,83	12,25	14,42	15,08	15,67	15,83	14,50	11,50	12,83	12,92	12,42	13,33	11,75	12,92	11,67	11,50	14,08
	sd	2,92	2,64	2,80	2,98	3,41	2,81	3,48	3,68	4,20	2,91	3,71	3,30	3,06	3,85	3,50	2,38	2,64	3,70	4,38	3,12
	t-value	0,396	-0,764	1,546	0,388	-1,268	1,130	1,578	2,042	1,926	1,191	-1,870	-0,700	-0,661	-0,975	-0,165	-2,548	-0,764	-1,716	-1,582	0,648
	p-value	0,700	0,461	0,150	0,706	0,231	0,283	0,143	0,066	0,080	0,259	0,088	0,498	0,522	0,350	0,872	**0,027**	0,461	0,114	0,142	0,530
5	mean	13,83	14,17	13,83	13,25	13,58	13,50	13,58	13,25	11,42	12,42	14,00	12,83	11,50	13,33	13,00	12,58	11,92	13,92	12,58	13,25
	sd	1,64	2,52	1,53	2,30	1,62	1,31	1,08	1,71	1,78	2,19	2,26	2,12	1,17	1,67	1,35	1,62	1,51	2,57	1,62	1,60
	t-value	0,703	0,918	0,756	-0,376	0,178	0,000	0,266	-0,506	-4,051	-1,711	0,768	-1,087	-5,933	-0,346	-1,285	-1,959	-3,644	0,561	-1,959	-0,540
	p-value	0,497	0,378	0,466	0,714	0,862	1,000	0,795	0,623	**0,002**	0,115	0,459	0,300	**0,000**	0,736	0,225	0,076	**0,004**	0,586	0,076	0,600
6	mean	13,25	13,33	14,75	13,17	12,58	11,67	13,50	13,58	13,75	13,50	12,75	13,00	13,00	13,33	14,08	13,75	13,00	14,08	13,25	12,17
	sd	1,06	1,44	2,56	2,48	1,31	2,42	1,45	1,93	2,99	1,51	3,33	2,56	1,65	2,06	1,68	2,22	1,86	1,98	1,14	1,64
	t-value	-0,821	-0,402	1,690	-0,466	-2,421	-2,619	0,000	0,150	0,290	0,000	-0,779	-0,677	-1,049	-0,280	1,205	0,390	-0,932	1,023	-0,761	-2,812
	p-value	0,429	0,695	0,119	0,651	**0,034**	**0,024**	1,000	0,884	0,777	1,000	0,452	0,512	0,317	0,784	0,253	0,704	0,371	0,328	0,463	**0,017**
7	mean	14,17	13,92	13,83	14,25	11,42	11,83	13,00	14,17	13,00	14,92	11,50	13,67	13,83	11,83	14,42	15,00	13,50	14,17	12,92	12,42
	sd	1,40	1,00	1,99	1,54	1,44	1,19	1,28	2,12	1,91	1,44	2,75	1,67	1,40	1,85	1,62	1,86	1,45	1,34	0,90	1,73
	t-value	1,646	1,449	0,580	1,682	-5,000	-4,838	-1,354	1,087	-0,908	3,400	-2,522	0,346	0,823	-3,120	1,959	2,796	0,000	1,727	-2,244	-2,169
	p-value	0,128	0,175	0,574	0,121	**0,000**	**0,001**	0,203	0,300	0,383	**0,006**	**0,028**	0,736	0,428	**0,010**	0,076	**0,017**	1,000	0,112	**0,046**	0,053

Table 3. Continuation

VP	image pair	AB	AC	AD	AE	BC	BD	BE	CD	CE	DE	VW	VX	VY	VZ	WX	WY	WZ	XY	XZ	YZ
8	**mean**	13,33	12,58	13,00	14,00	14,33	13,50	15,25	13,58	14,08	14,75	11,58	12,33	14,33	15,00	14,08	12,50	14,00	12,67	13,92	14,83
	sd	1,83	2,57	1,86	2,45	2,61	1,83	2,86	2,78	2,91	3,28	2,81	2,64	3,03	3,59	3,00	2,02	3,81	3,42	1,73	1,95
	t-value	-0,316	-1,233	-0,932	0,707	1,108	0,000	2,116	0,104	0,695	1,321	-2,362	-1,531	0,954	1,446	0,674	-1,713	0,454	-0,844	0,834	2,373
	p-value	0,758	0,243	0,371	0,494	0,291	1,000	0,058	0,919	0,501	0,213	**0,038**	0,154	0,360	0,176	0,514	0,115	0,659	0,417	0,422	**0,037**
9	**mean**	13,92	13,75	14,17	14,25	13,92	13,33	14,67	12,75	13,92	15,92	12,92	12,50	13,50	14,50	14,42	13,50	15,08	14,00	13,92	14,83
	sd	1,24	1,06	1,99	1,60	1,56	2,15	1,37	1,22	1,38	1,68	1,78	1,88	1,09	1,38	1,44	1,24	1,38	1,13	1,08	1,40
	t-value	1,164	0,821	1,159	1,621	0,923	-0,269	2,948	-2,138	1,047	4,994	-1,134	-1,840	0,000	2,507	2,200	0,000	3,978	1,535	1,332	3,291
	p-value	0,269	0,429	0,271	0,133	0,376	0,793	**0,013**	0,056	0,318	**0,000**	0,281	0,093	1,000	**0,029**	0,050	1,000	**0,002**	0,153	0,210	**0,007**
10	**mean**	13,50	13,00	12,08	12,17	11,83	12,25	13,50	14,75	13,83	14,75	10,00	13,17	14,58	12,92	15,75	14,92	13,08	14,67	11,92	11,75
	sd	1,51	1,76	1,31	1,70	1,19	1,71	1,78	1,42	3,04	1,96	1,13	1,19	2,15	2,23	2,01	2,11	1,38	2,77	1,78	1,42
	t-value	0,000	-0,985	-3,742	-2,722	-4,838	-2,529	0,000	3,045	0,380	2,209	-10,747	-0,968	1,744	-0,904	3,886	2,327	-1,047	1,457	-3,079	-4,262
	p-value	1,000	0,346	**0,003**	**0,020**	**0,001**	**0,028**	1,000	**0,011**	0,711	**0,049**	**0,000**	0,354	0,109	0,385	**0,003**	**0,040**	0,318	0,173	**0,010**	**0,001**
11	**mean**	14,08	13,25	14,50	12,92	13,50	14,00	13,83	12,50	12,33	12,67	13,17	12,00	14,25	14,75	14,17	15,17	13,83	14,92	13,33	13,50
	sd	1,93	1,42	1,68	1,44	1,73	1,28	1,19	1,73	2,23	1,37	1,53	1,71	2,09	1,76	1,70	1,59	1,47	2,31	1,30	1,51
	t-value	1,048	-0,609	2,064	-1,400	0,000	1,354	0,968	-2,000	-1,813	-2,106	-0,756	-3,047	1,241	2,454	1,361	3,640	0,787	2,120	-0,443	0,000
	p-value	0,317	0,555	0,063	0,189	1,000	0,203	0,354	0,071	0,097	0,059	0,466	**0,011**	0,241	**0,032**	0,201	**0,004**	0,448	0,058	0,666	1,000
12	**mean**	13,25	14,58	13,42	13,33	13,75	13,83	13,42	12,92	13,83	15,00	11,08	12,42	12,92	12,33	14,08	14,17	13,67	14,67	11,42	11,75
	sd	1,96	2,15	2,54	1,56	1,86	2,41	2,15	2,02	1,99	1,91	2,97	1,38	1,44	1,92	1,08	2,08	1,56	1,50	1,31	1,29
	t-value	-0,442	1,744	-0,114	-0,371	0,464	0,480	-0,134	-1,000	0,580	2,725	-2,820	-2,721	-1,400	-2,102	1,865	1,109	0,371	2,699	-5,503	-4,706
	p-value	0,667	0,109	0,912	0,718	0,651	0,641	0,896	0,339	0,574	**0,020**	**0,017**	**0,020**	0,189	0,059	0,089	0,291	0,718	**0,021**	**0,000**	**0,001**
13	**mean**	13,92	14,83	14,83	13,75	14,33	13,00	14,00	13,75	13,08	14,08	13,42	12,17	11,92	11,50	13,25	12,92	11,92	11,67	12,25	12,67
	sd	1,38	2,55	1,53	1,60	1,78	1,86	1,48	2,86	2,57	1,24	2,02	2,17	1,83	1,31	1,71	1,44	1,44	2,31	1,36	1,61
	t-value	1,047	1,810	3,024	0,540	1,626	-0,932	1,173	0,302	-0,561	1,629	-0,143	-2,131	-2,994	-5,272	-0,506	-1,400	-3,800	-2,750	-3,191	-1,788
	p-value	0,318	0,098	**0,012**	0,600	0,132	0,371	0,266	0,768	0,586	0,131	0,889	0,056	**0,012**	**0,000**	0,623	0,189	**0,003**	**0,019**	**0,009**	0,101
14	**mean**	13,42	13,83	14,00	13,67	14,17	14,08	13,25	14,58	12,42	13,92	12,58	14,25	13,25	11,08	16,42	12,92	12,67	13,42	13,17	14,25
	sd	3,00	2,86	1,86	4,68	3,95	3,50	3,52	5,42	2,84	4,06	2,57	2,83	2,96	3,18	3,90	3,09	3,39	3,18	3,46	3,60
	t-value	-0,096	0,404	0,932	0,123	0,585	0,577	-0,246	0,693	-1,320	0,356	-1,233	0,917	-0,293	-2,636	2,594	-0,654	-0,851	-0,091	-0,334	0,722
	p-value	0,925	0,694	0,371	0,904	0,571	0,576	0,810	0,503	0,214	0,729	0,243	0,379	0,775	**0,023**	**0,025**	0,526	0,413	0,929	0,745	0,485

Table 3. Continuation

VP	image pair	AB	AC	AD	AE	BC	BD	BE	CD	CE	DE	VW	VX	VY	VZ	WX	WY	WZ	XY	XZ	YZ
15	mean	14,17	13,83	13,50	13,33	12,17	12,17	13,58	14,83	12,83	13,50	13,08	14,25	14,33	13,42	16,25	14,67	14,42	12,58	12,08	13,92
	sd	2,12	0,83	1,83	1,37	1,34	1,27	1,73	2,08	2,55	1,62	1,56	1,36	1,56	1,98	1,71	1,07	1,44	2,23	2,07	1,16
	t-value	1,087	1,383	0,000	-0,421	-3,454	-3,645	0,167	2,219	-0,905	0,000	-0,923	1,915	1,854	-0,146	5,564	3,766	2,200	-1,421	-2,376	1,239
	p-value	0,300	0,194	1,000	0,682	**0,005**	**0,004**	0,870	**0,048**	0,385	1,000	0,376	0,082	0,091	0,886	**0,000**	**0,003**	0,050	0,183	**0,037**	0,241
16	mean	13,33	13,50	14,83	13,50	13,50	12,50	14,00	13,17	13,42	12,67	13,83	13,33	13,08	13,17	13,33	13,83	13,83	14,83	13,50	13,58
	sd	1,23	0,80	4,02	2,47	1,62	1,93	1,71	1,90	1,83	1,83	1,34	1,61	1,56	1,03	1,83	1,40	1,85	1,34	1,38	1,38
	t-value	-0,469	0,000	1,149	0,000	0,000	-1,794	1,016	-0,608	-0,158	-1,581	0,864	-0,358	-0,923	-1,121	-0,316	0,823	0,624	3,454	0,000	0,209
	p-value	0,648	1,000	0,275	1,000	1,000	0,100	0,332	0,555	0,878	0,142	0,406	0,727	0,376	0,286	0,758	0,428	0,545	**0,005**	1,000	0,838
17	mean	13,92	14,75	13,08	13,67	13,25	13,25	13,75	12,75	14,58	14,58	12,58	14,33	12,83	13,75	15,33	12,67	14,17	13,92	12,33	13,42
	sd	2,39	3,33	2,27	2,71	2,30	2,67	1,42	2,86	4,29	2,54	2,64	3,39	2,41	2,38	2,35	2,35	2,52	3,06	1,97	1,00
	t-value	0,604	1,299	-0,635	0,213	-0,376	-0,325	0,609	-0,907	0,874	1,478	-1,201	0,851	-0,960	0,364	2,704	-1,229	0,918	0,472	-2,052	-0,290
	p-value	0,558	0,221	0,539	0,835	0,714	0,751	0,555	0,384	0,401	0,167	0,255	0,413	0,358	0,723	**0,020**	0,245	0,378	0,646	0,065	0,777
18	mean	13,75	13,33	13,25	13,83	12,92	13,42	13,50	14,42	15,75	15,42	11,83	13,50	13,83	13,17	16,42	15,00	13,92	14,00	11,83	13,25
	sd	1,42	2,10	1,76	1,85	1,51	1,24	1,83	2,35	1,14	2,19	1,80	1,78	0,94	1,03	2,23	0,95	1,24	2,13	2,04	1,48
	t-value	0,609	-0,274	-0,491	0,624	-1,343	-0,233	0,000	1,349	6,848	3,027	-3,206	0,000	1,232	-1,121	4,522	5,450	1,164	0,812	-2,834	-0,583
	p-value	0,555	0,789	0,633	0,545	0,206	0,820	1,000	0,204	**0,000**	**0,012**	**0,008**	1,000	0,244	0,286	**0,001**	**0,000**	0,269	0,434	**0,016**	0,571
19	mean	13,33	11,58	12,17	13,50	12,33	14,58	13,75	13,67	12,42	15,00	11,00	12,42	13,42	13,33	14,08	10,75	11,92	11,50	14,42	14,00
	sd	3,28	2,23	2,89	3,18	2,53	2,43	2,99	3,42	2,91	2,56	2,73	2,35	2,02	2,42	3,18	2,14	3,18	3,40	3,26	1,81
	t-value	-0,176	-2,972	-1,600	0,000	-1,595	1,545	0,290	0,169	-1,291	2,031	-3,172	-1,595	-0,143	-0,238	0,636	-4,457	-1,727	-2,039	0,974	0,957
	p-value	0,864	**0,013**	0,138	1,000	0,139	0,151	0,777	0,869	0,223	0,067	**0,009**	0,139	0,889	0,816	0,538	**0,001**	0,112	0,066	0,351	0,359
20	mean	13,08	13,50	11,75	11,75	14,67	11,50	12,25	13,00	10,67	14,33	10,25	11,17	13,25	11,58	14,75	13,83	13,25	14,08	14,50	14,75
	sd	2,27	3,61	3,31	4,20	3,11	3,71	2,60	4,18	2,87	2,57	3,28	2,17	3,25	1,98	2,42	2,37	2,14	3,42	2,15	2,34
	t-value	-0,635	0,000	-1,834	-1,443	1,298	-1,870	-1,667	-0,415	-3,419	1,123	-3,434	-3,730	-0,266	-3,361	1,792	0,488	-0,405	0,590	1,609	1,850
	p-value	0,539	1,000	0,094	0,177	0,221	0,088	0,124	0,686	**0,006**	0,285	**0,006**	**0,003**	0,795	**0,006**	0,101	0,635	0,693	0,567	0,136	0,091

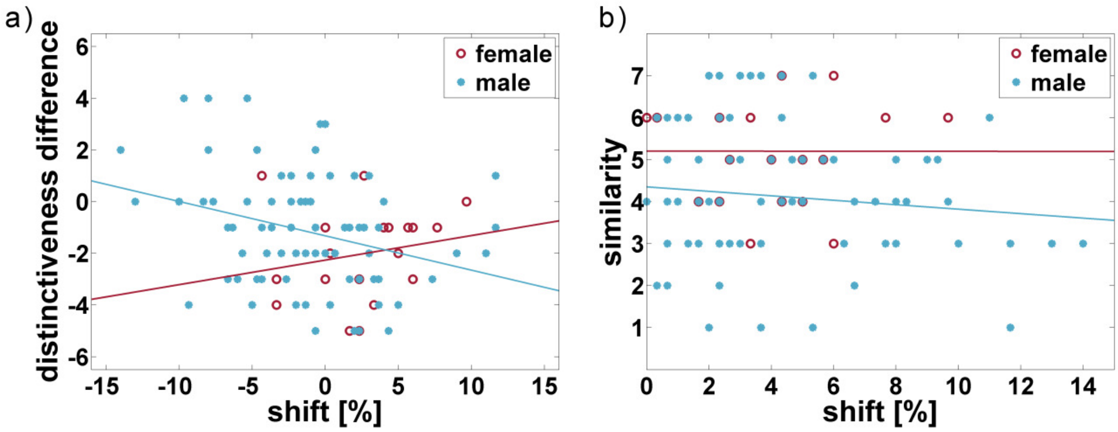

Figure 51. Correlations between significant shifts and distinctiveness and similarity. a) Distinctiveness differences were highly correlated to the shifts in morph-level for male faces towards the more distinctive face (negative y-axis value). For female faces no correlation was present. b) Similarity ratings were uncorrelated from shifts in morph-level for both female and male faces. Each marker indicates one significant rating per face pair and participant. The x-axes show shifts in morph-level, the y-axes show the difference in distinctiveness ratings (a) and similarity ratings (b).

Experiment 3 - Primed face recognition

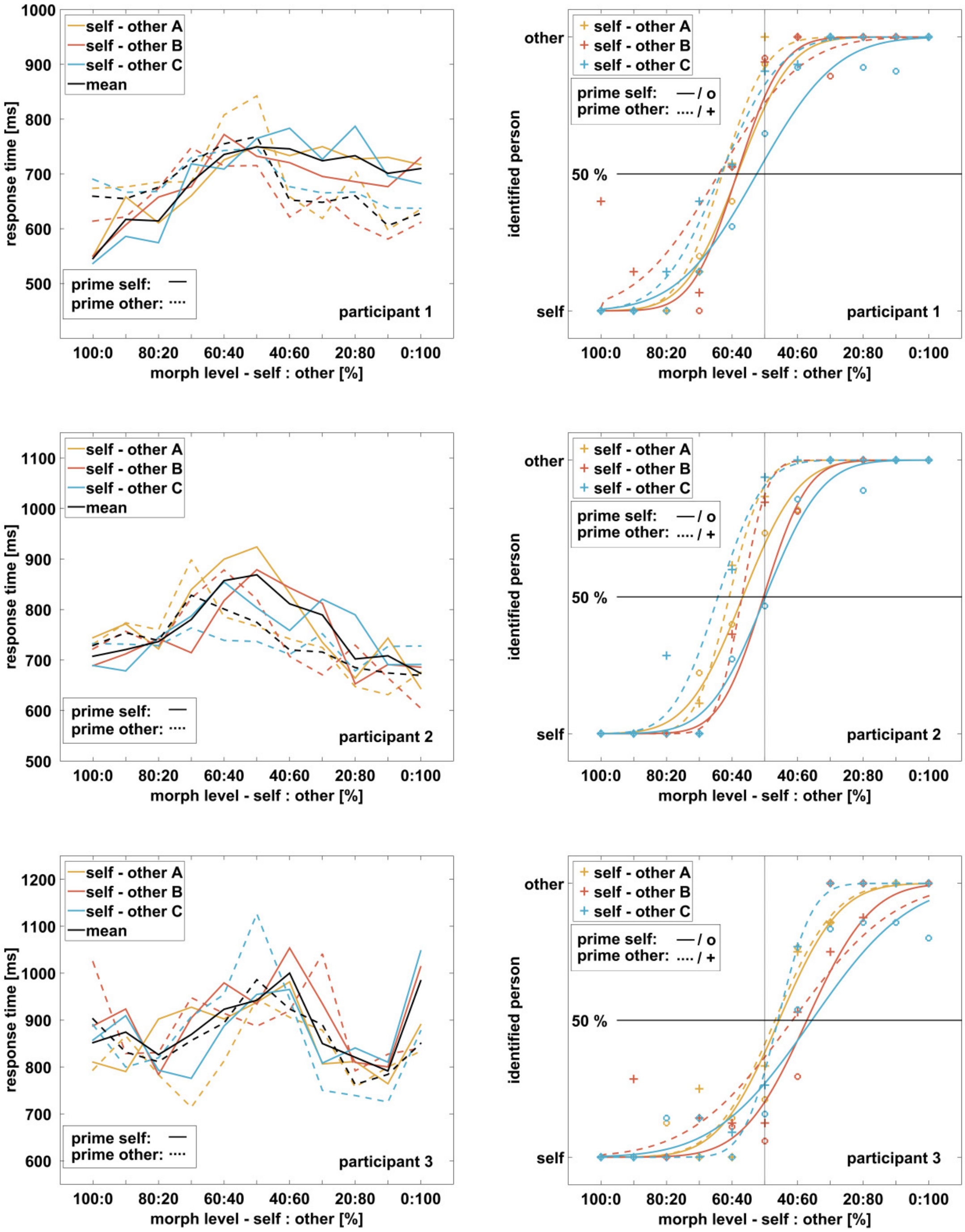

Figure 52. Response time courses and identified person. Response times (left side) of participants 1 to 3 and their respective answering behavior (right side) are shown. Response times typically increased around the 50 % morph-level as the test image became more ambiguous. Primes that showed the participants themselves (solid lines) led to more identification as "self" in the pictures. Solid lines and circle markers indicate primes on "self" while dotted lines and plus markers indicate primes on "other". Colors indicate different face pairs for testing. Markers (+ and o) show the average answers at the given morph-level with psychometric curves fitted to these data. X-axes display morph-levels in percent and y-axes response times in milliseconds or identified person.

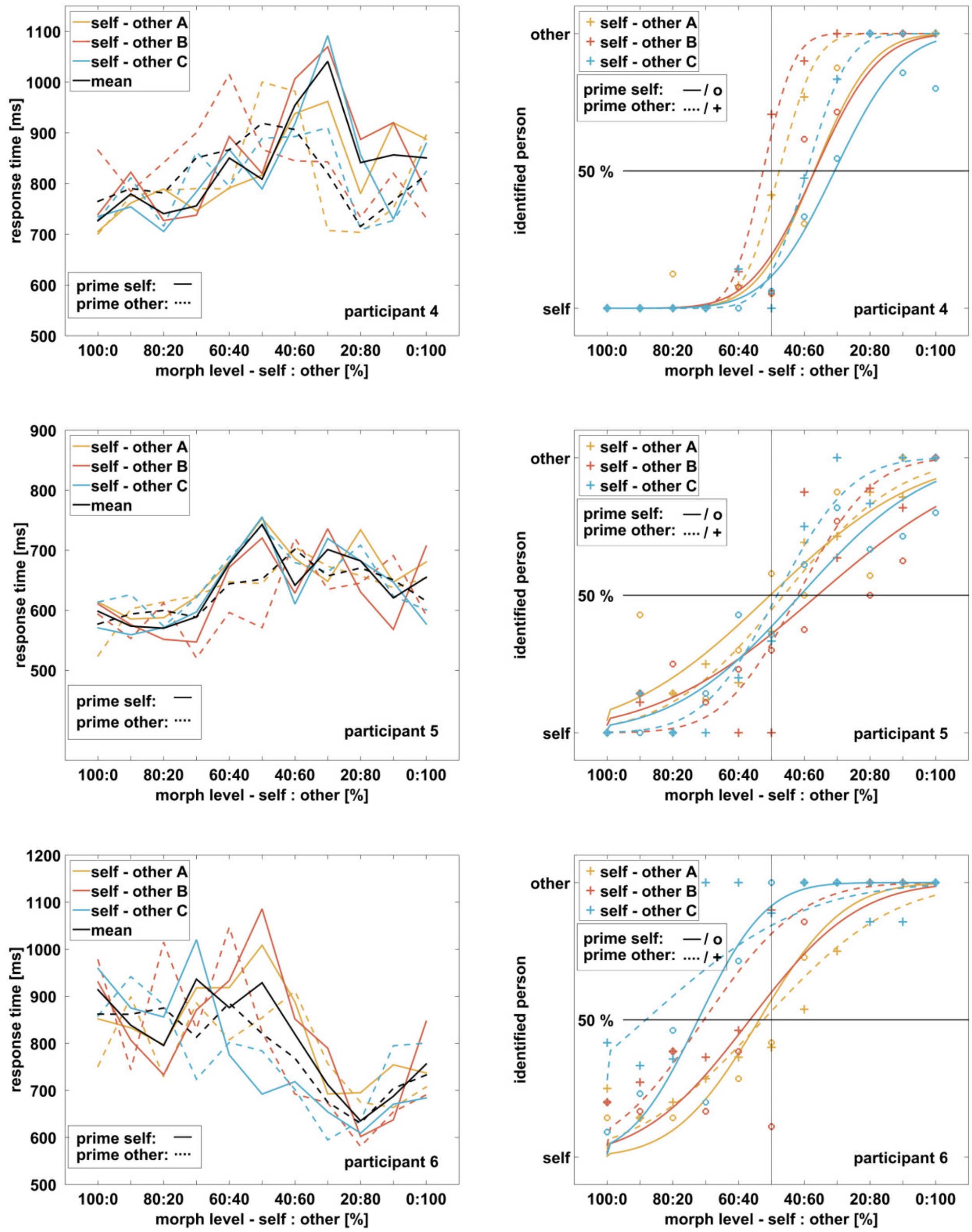

Figure 53. Response time courses and identified person. Response times (left side) of participants 4 to 6 and their respective answering behavior (right side) are shown. Response times typically increased around the 50 % morph-level as the test image became more ambiguous. Primes that showed the participants themselves (solid lines) led to more identification as "self" in the pictures. Solid lines and circle markers indicate primes on "self" while dotted lines and plus markers indicate primes on "other". Colors indicate different face pairs for testing. Markers (+ and o) show the average answers at the given morph-level with psychometric curves fitted to these data. X-axes display morph-levels in percent and y-axes response times in milliseconds or identified person.

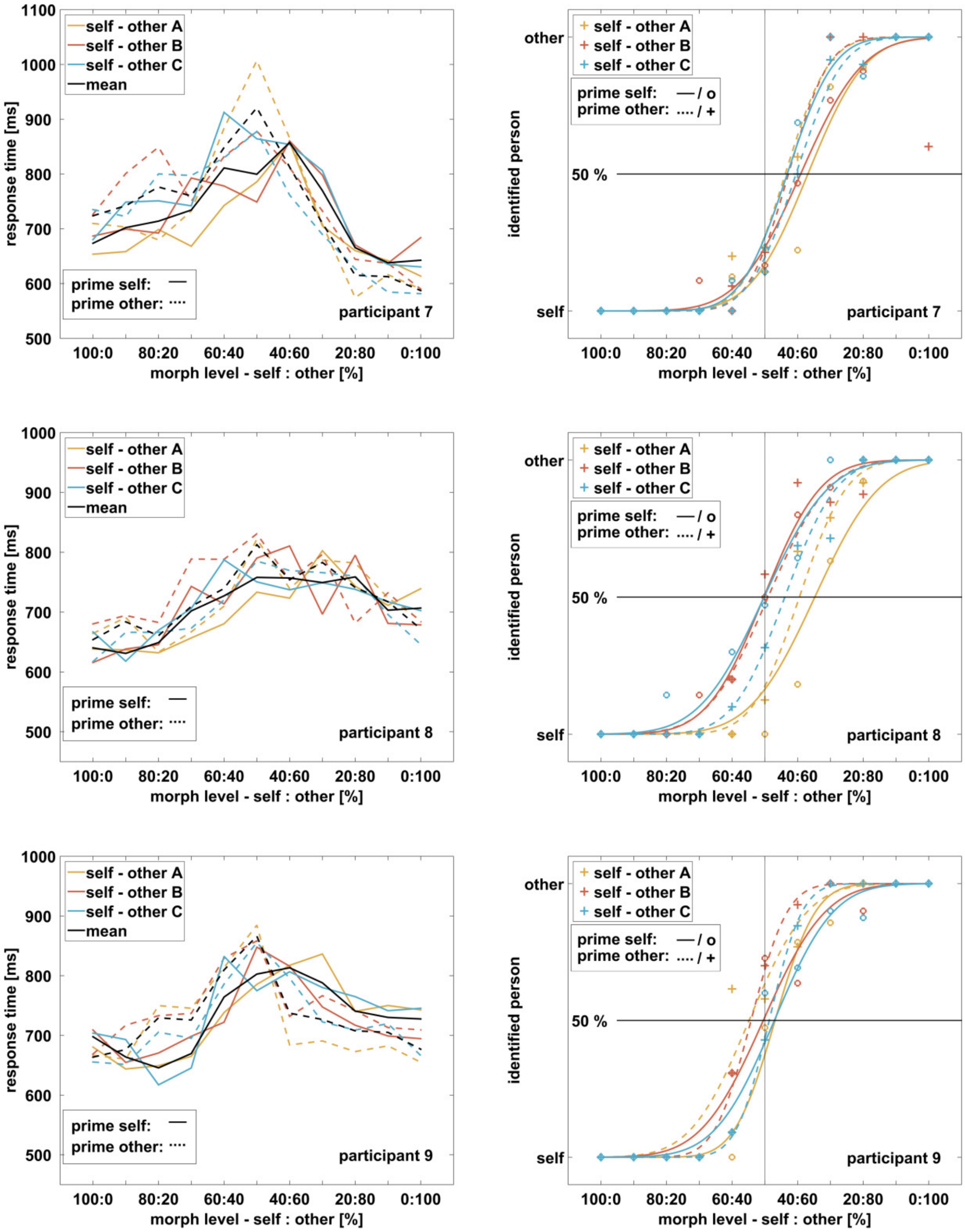

Figure 54. Response time courses and identified person. Response times (left side) of participants 7 to 9 and their respective answering behavior (right side) are shown. Response times typically increased around the 50 % morph-level as the test image became more ambiguous. Primes that showed the participants themselves (solid lines) led to more identification as "self" in the pictures. Solid lines and circle markers indicate primes on "self" while dotted lines and plus markers indicate primes on "other". Colors indicate different face pairs for testing. Markers (+ and o) show the average answers at the given morph-level with psychometric curves fitted to these data. X-axes display morph-levels in percent and y-axes response times in milliseconds or identified person.

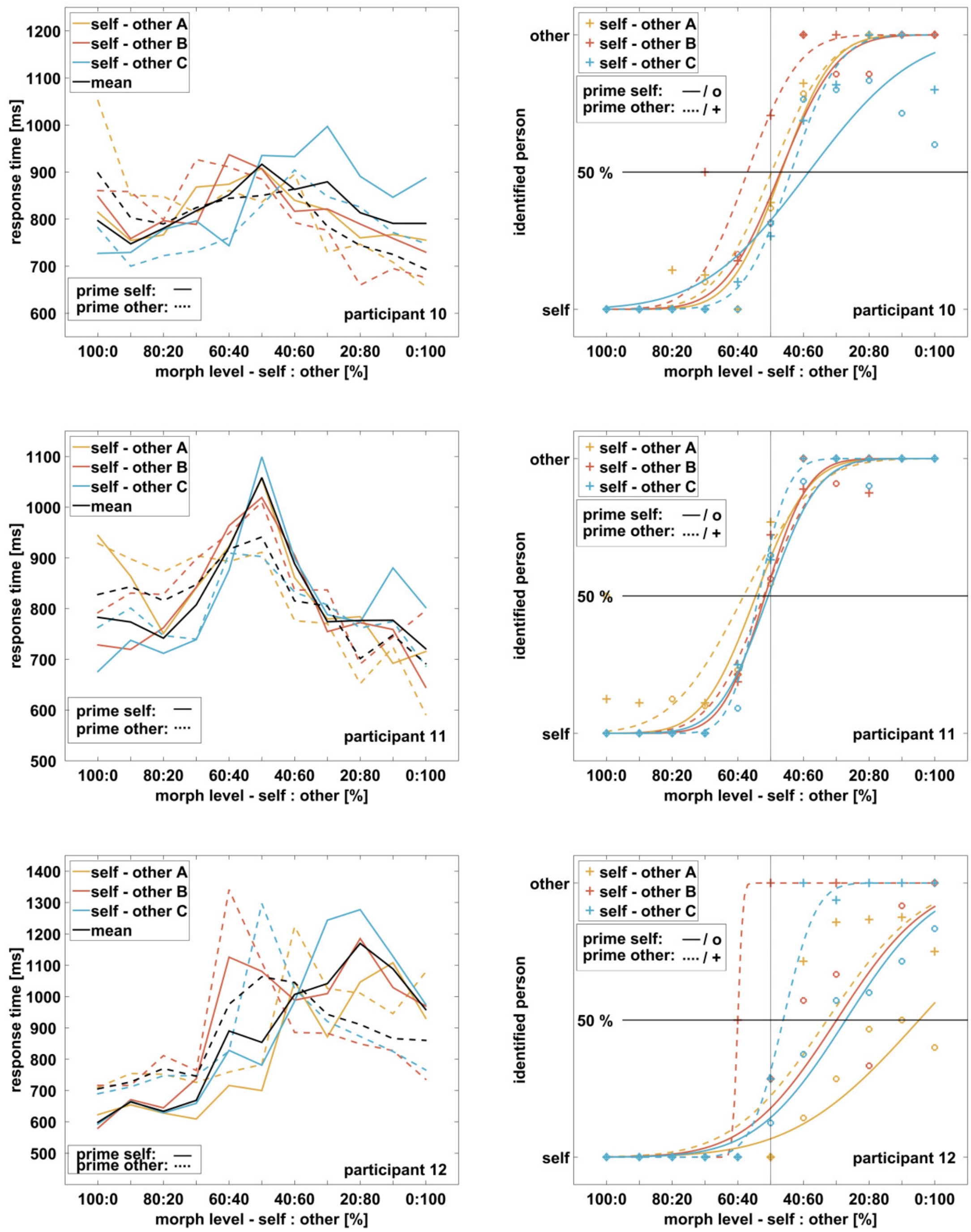

Figure 55. Response time courses and identified person. Response times (left side) of participants 10 to 12 and their respective answering behavior (right side) are shown. Response times typically increased around the 50 % morph-level as the test image became more ambiguous. Primes that showed the participants themselves (solid lines) led to more identification as "self" in the pictures. Solid lines and circle markers indicate primes on "self" while dotted lines and plus markers indicate primes on "other". Colors indicate different face pairs for testing. Markers (+ and o) show the average answers at the given morph-level with psychometric curves fitted to these data. X-axes display morph-levels in percent and y-axes response times in milliseconds or identified person.

Experiment 4 - View-based spatial memory

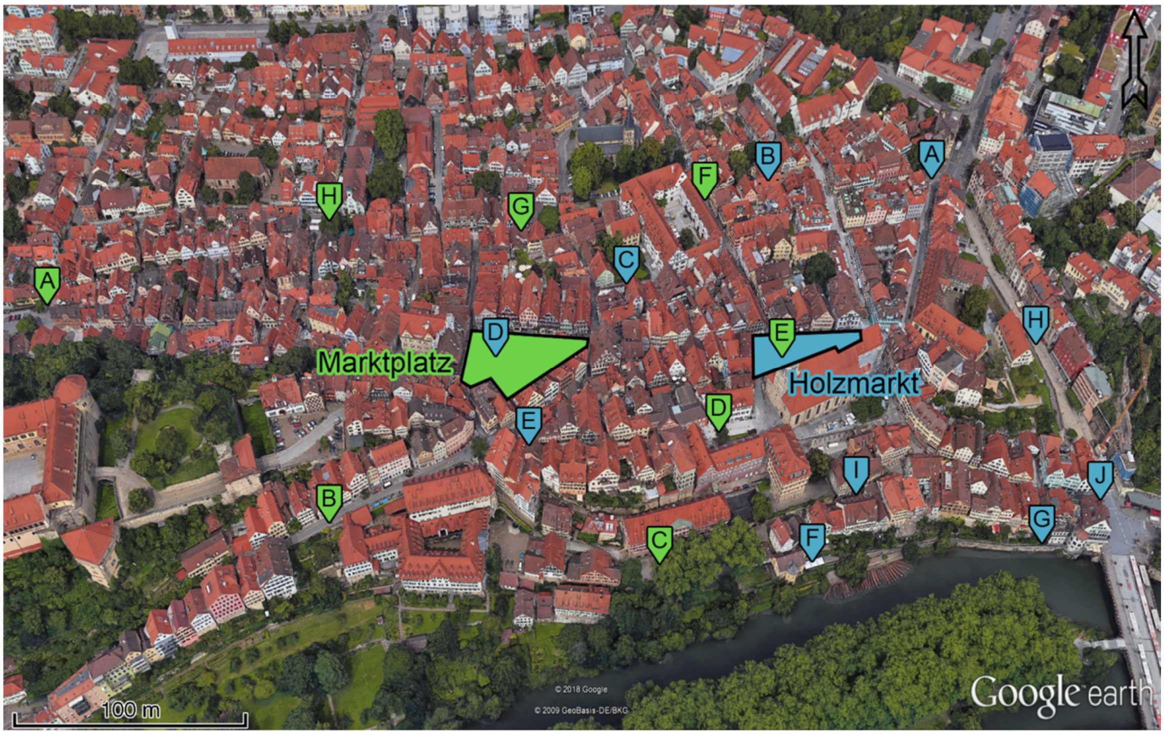

Figure 56. Aerial photograph of downtown Tübingen. Displayed are the outlines of the target places "Holzmarkt" (timber market, blue) and "Marktplatz" (market place, green) as well as their respective interview locations (letters A to J). At the bottom area, the river can be seen with the two interview locations F and G right next to it. Map sources: Google Earth (Google LLC, USA) & GeoBasis-DE/BKG (Bundesamt für Kartographie und Geodäsie, Germany).

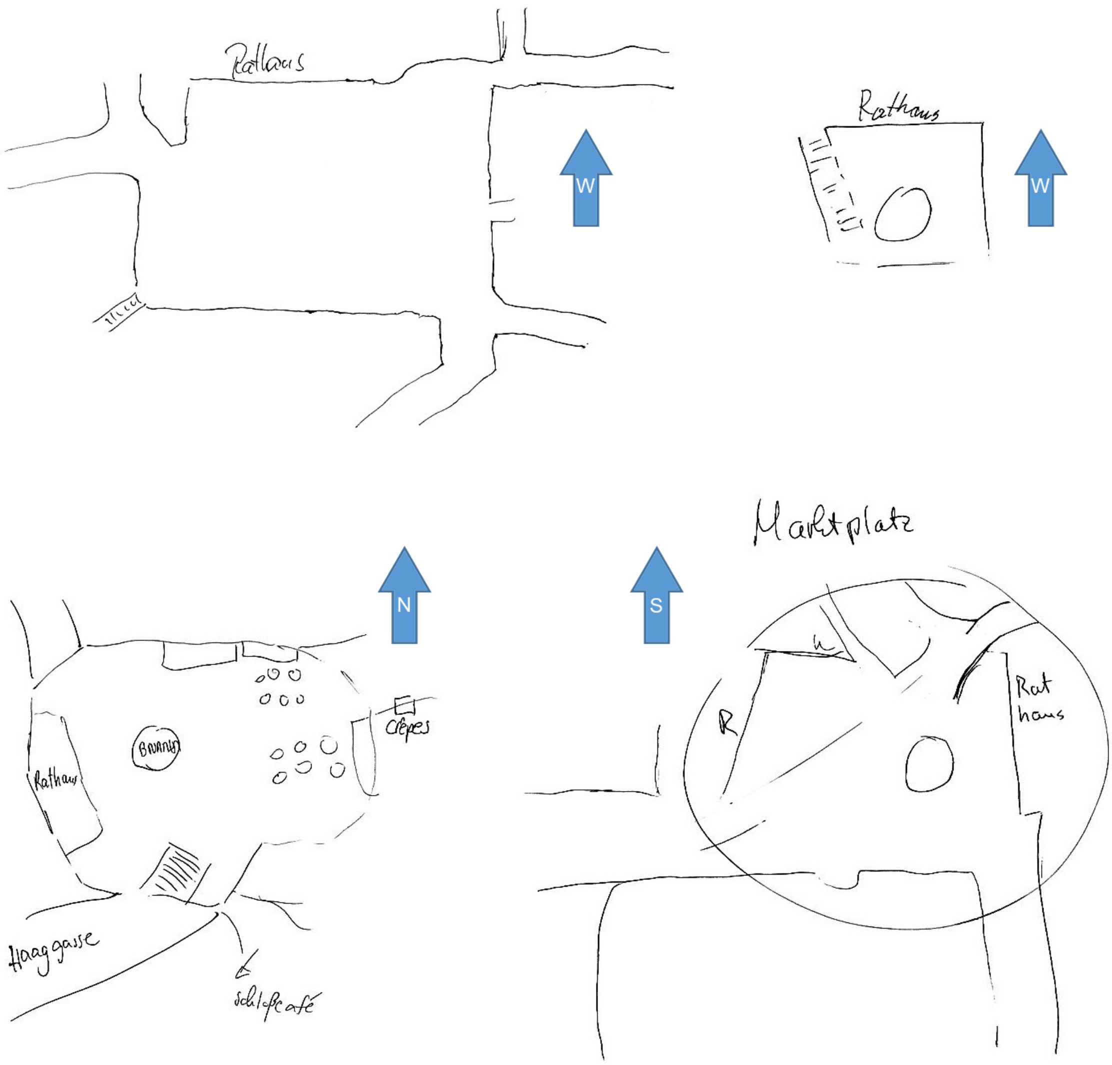

Figure 57. Examples of sketches of the "Marktplatz" from four participants. The blue arrows indicate the orientation the sketches were rated in by the three raters during analysis. Note the inscription "Rathaus" in all four sketches, referring to the landmark town hall that is prominently located at the western side of this square. The circles mark a fountain in the center of the square. The "parallel" lines mark flights of stairs. There are two stairs located at two corners of the square. Sketches two and three depict the bigger one, while in sketch one the smaller stairs was drawn. In the third sketch, also other landmarks and streets have been named. In the fourth sketch, the participant also marked and labeled the area the market place covers.

Experiment 5 - View recognition of places

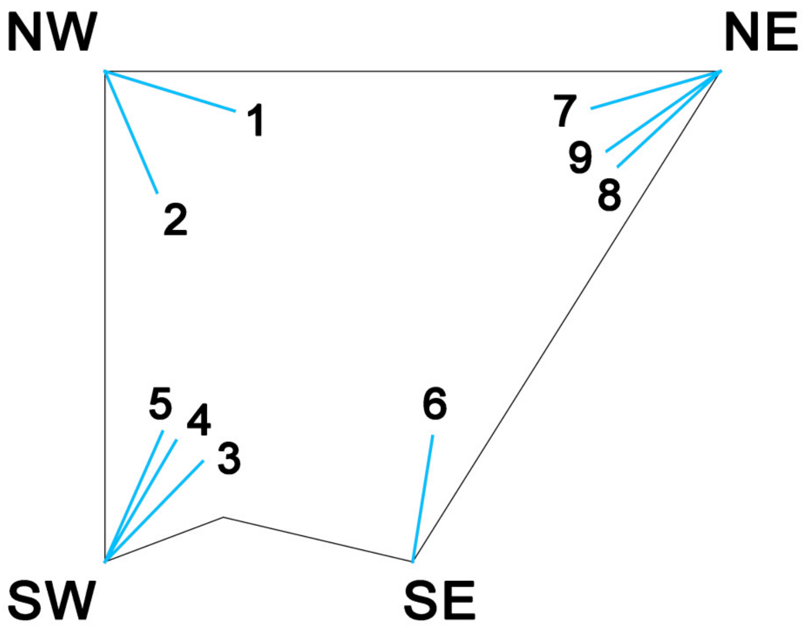

Figure 58. Locations and orientations of the market place pictures. Depicted are the viewing directions (blue) and numbers of the nine pictures taken at the four entrances (NW, NE, SE, SW) to the market place (dark outline). Please also see Figure 59 for the photographs.

Figure 59. Photographs of the market place. The nine images were taken at the four entrances to the market place and were shown to the participants; please also see Figure 58 for the orientations. During the experiment, the numbers were not visible.

Table 4. Participants' signal detection performance per interview location. The participants' hit and correct rejection rates on detecting market place pictures were always above 90 %, leading to sensitivity indices (d') of 3.6 and better for all interview locations.

interview location	A	B	C	D	E	F	M
hit rate [%]	99.4	99.1	98.1	98.3	98.1	99.4	97.8
correct rejection rate [%]	93.0	92.0	94.6	96.8	94.4	95.2	97.4
sensitivity index (d')	4.012	3.755	3.694	3.975	3.659	4.202	3.954

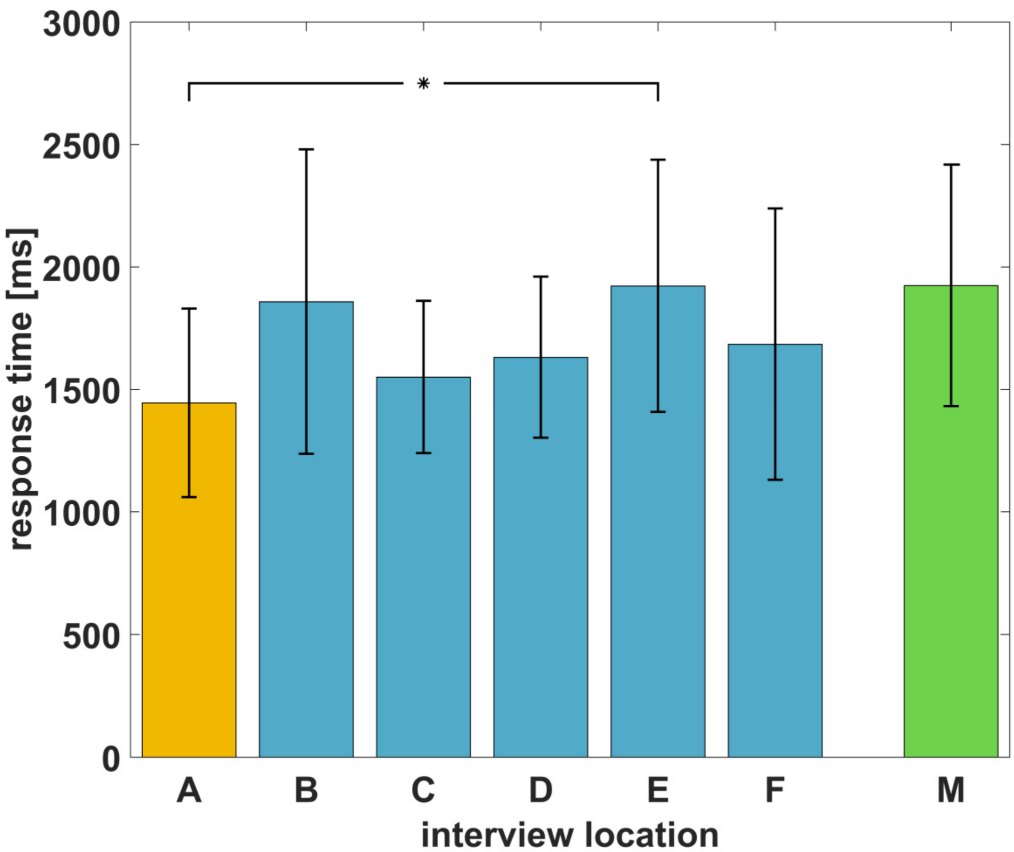

Figure 60. Average response time per interview location for all photographs. Response times were significantly different for the interview locations. The y-axis shows the response time, the x-axis the interview locations. Error bars show the standard deviation.

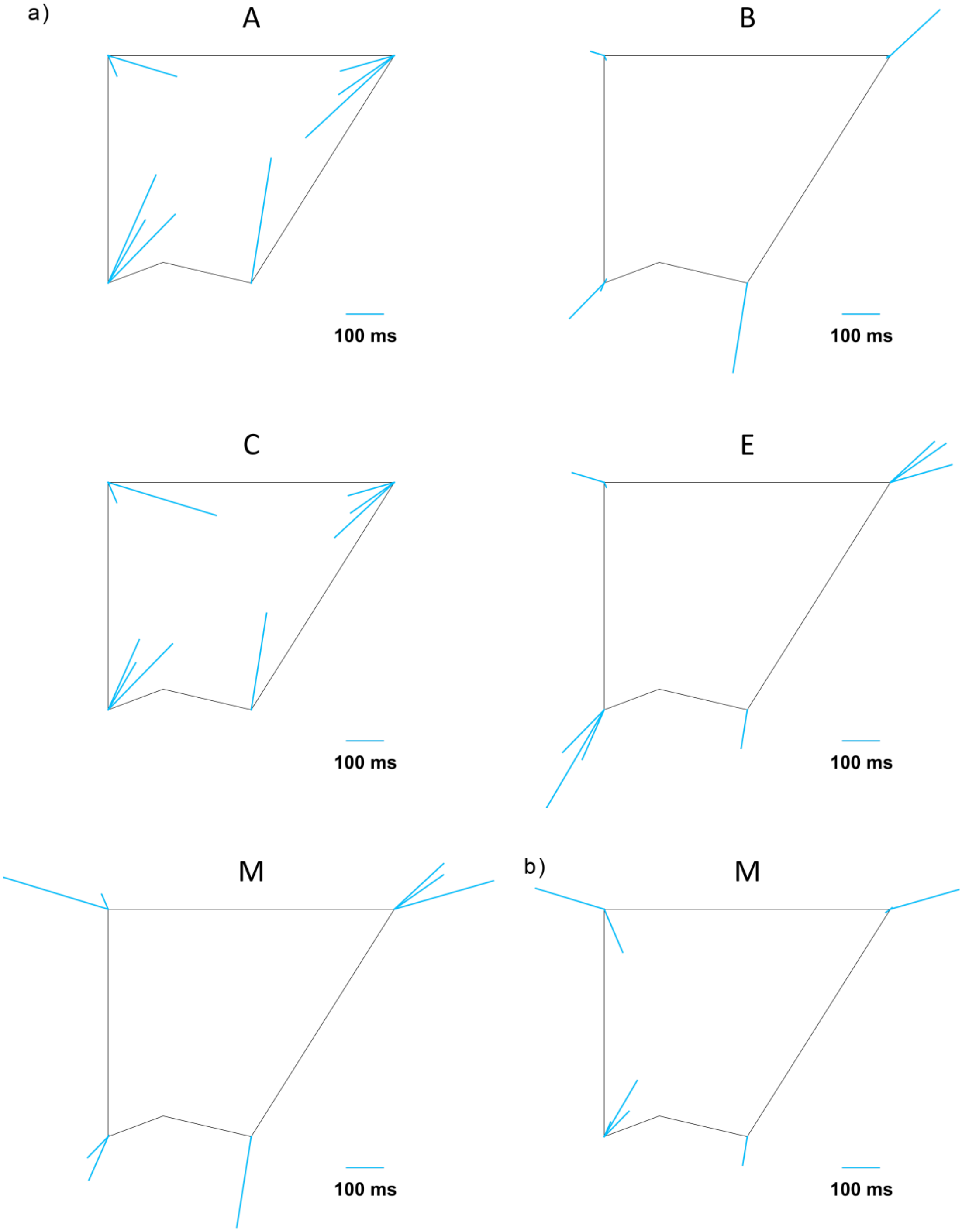

Figure 61. Response time differences at interview locations. a) Response time differences after subtraction of the total mean are shown. Blue lines indicate view point and angle of the presented photographs of the target place and their length response time differences. Lines pointing inwards indicate smaller response times, lines pointing outwards show responses that took longer than average. The dark outlines sketch the market place. For interview locations D and F, see results section of experiment 5 [Figure 44 b)]. b) Response time differences after subtraction of the local mean for interview location M (market place) are shown.

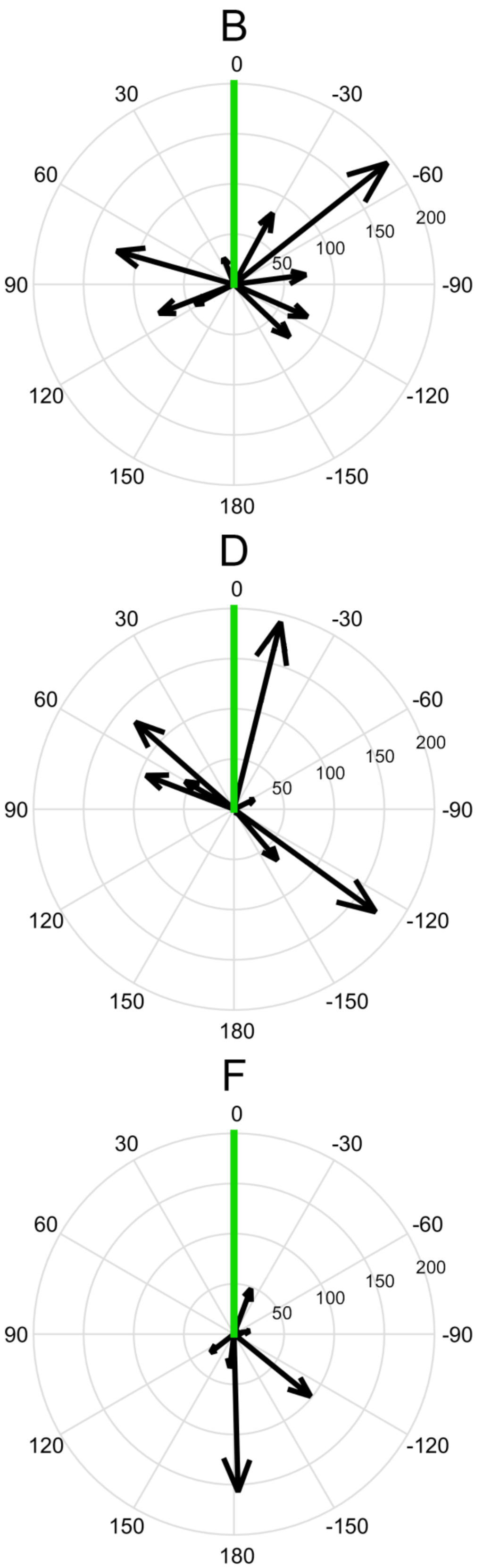

Figure 62. Response time vectors of near interview locations (B, C, D, E, F). The green line indicates the air-line direction from the interview location to the four entrances on the market place each photograph was taken at, the black arrows show the response time vectors for the nine images of the market place. The longer the arrow, the stronger was the deviation from the average response time. Please note the different scales in ms (radius).

General discussion

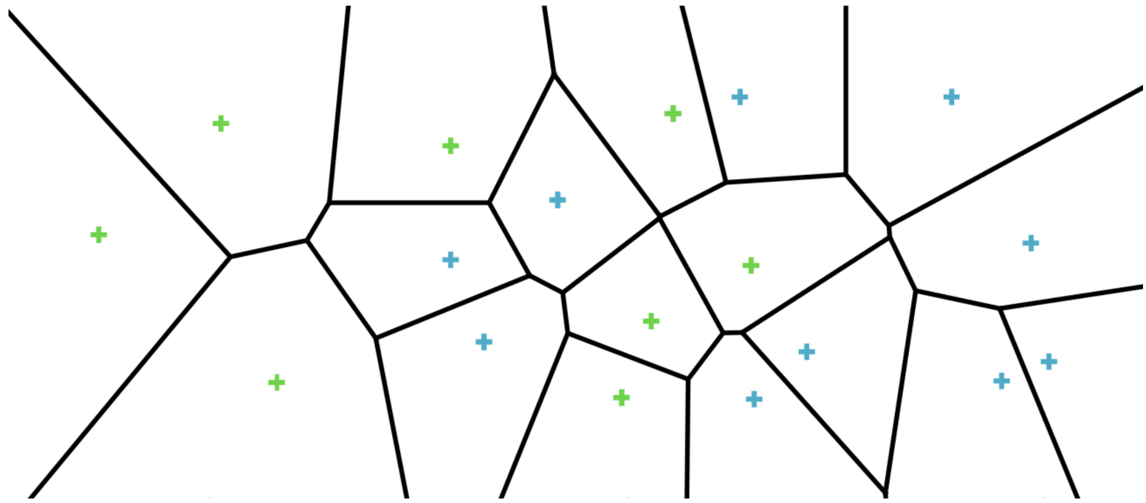

Figure 63. Example of a two-dimensional Voronoi diagram. Such a Voronoi diagram can also be created for high-dimensional spaces like face-space. There, each marker represents a face with its attributes that span the face-space. The black lines mark the borders of the cells, separating each data point equidistantly from the neighboring ones. A cell is thought to be the capture area for that face, which means variations of the visual stimulus are still attributed to that face. For this example diagram, the coordinates of the interview locations (blue and green markers) of experiment 4 were used. Please note that in the example here with the interview locations, such a capture area does not exist, because the interview locations are supposed to have no meaning to the participants, and therefore they should not be not represented in memory.

13. Contributions

All experimental designs as well as the results of this doctoral thesis were designed and discussed together with Prof. Dr. Hanspeter A. Mallot.

In experiment 1, all data analyses, figures and written parts were done by me. Experimental data was collected under my supervision by Alisa Volkert who also used the data for her bachelor thesis. The programs for data collection and data analysis were written by me. No parts of the bachelor thesis were used for this thesis.

In experiment 2, all data analyses, figures and written parts were done by me. Experimental data was collected under my supervision by Esther Kutter who also used the data for her bachelor thesis. The program for data collection was written by Esther Kutter; the program for data analysis was written by me. No parts of the bachelor thesis were used for this thesis.

In experiment 3, all data analyses, figures and written parts were done by me. Experimental data for experiment 3 a) was collected by me and for experiment 3 b) it was collected under my supervision by Frank Riemer who also used the data for his bachelor thesis. The programs for data collection and data analysis were written by me. No parts of the bachelor thesis were used for this thesis.

Experiment 4 has been published as an original research paper "View-Based Organization and Interplay of Spatial Working and Long-Term Memories" by Röhrich, Hardieß & Mallot in 2014 at PLOS ONE, 9(11), e112793 (https://doi.org/10.1371/journal.pone.0112793). The publication was written by me together with Prof. Dr. Hanspeter A. Mallot and PD Dr. Gregor Hardieß. The text and figures of this publication were used with minor changes and additions in this thesis.
The idea for this experiment was developed by me together with Prof. Dr. Hanspeter A. Mallot. Experimental data were collected under my supervision by Niklas Binder for experiment 4 a) who also used the data for his bachelor thesis and by Julia Mayer for experiment 4 b) who used the data for her state exam essay. No parts of the bachelor thesis or state exam essay were used for this doctoral thesis. The programs for data analysis were written by me. Data analysis and figures were done by me in dialogue with Prof. Dr. Hanspeter A. Mallot and PD Dr. Gregor Hardieß. Written parts were equally done by me, Prof. Dr. Hanspeter A. Mallot and PD Dr. Gregor Hardieß. The mathematical model was developed by Prof. Dr. Hanspeter A. Mallot.

In experiment 5, all data analyses, figures and written parts were done by me. Experimental data was collected under my supervision by Amelie Haug who also used the data for her bachelor thesis. The program for data collection was written by Dr. Marc Halfmann. The program for data analysis was written by me. No parts of the bachelor thesis were used for this thesis.