Color, Metaphor and Culture

Empirical Foundations for User Interface Design

Inaugural-Dissertation
zur Erlangung der Doktorwürde der
Fakultät für Humanwissenschaften der
Julius-Maximilians-Universität Würzburg
vorlegelegt von Diana Löffler aus Berlin

Würzburg
2017

Erstgutachter: Prof. Dr.-Ing. Jörn Hurtienne

Zweitgutachter: Prof. Dr. Eng. Takashi Toriizuka

Tag des Kolloquiums: 15.05.2017

Meinem Vater.

Abstract

Using color in user interface design is both art and science. Often, designers focus on aesthetic properties of color, but neglect that it also carries meaning and entails profound psychological consequences. Color psychology, filling this gap, is in its infancy, and lacks a theoretical approach that predicts and explains color-meaning associations shared by a large group of people in a large variety of contexts.

To amend this situation, this work develops *Conceptual Metaphor Theory of Color* (CMToC), which predicts and explains cross-cultural and experience-based semantic color associations. The theory is based on the idea from cognitive linguistics that the study of metaphorical language provides valuable insights into our mental models involving color. A discussion of three types of metaphors that cover associations with physical and abstract concepts in light of existing empirical evidence provides the basis for deriving empirical research questions.

The first research question addresses the use of color for conveying physical information like *weight* in user interfaces. The results of four online surveys involving a total of 295 German and Japanese participants show the relative impact of hue, saturation and brightness for associations with 16 physical properties. Two thirds of these color associations were correctly predicted by CMToC. Participants frequently matched physical properties to colors based on sensorimotor correspondences and participants of both cultures did not considerably vary in their performance.

The second research question addresses the use of color for conveying abstract information like *importance* in user interfaces. In one experimental study, a total of 75 German and Japanese participants validated color-to-abstract mappings in form of color population stereotypes like *important is dark*. The majority of these color associations (86%) were correctly predicted by CMToC. Again, participants of both cultures did not considerably vary in their performance.

The third research question addresses whether predicted color associations with physical and abstract information are processed automatically as a precondition for intuitive use. The results of three studies involving a total of 85 German and Japanese participants show on the example of *temperature* that color automatically influences the identification speed of related physical properties, but not vice versa. Color and abstract information were not automatically associated.

As a result of these studies it can be concluded that predictions of CMToC are cross-culturally valid for user interface design. Derived implicit associations with physical properties and explicit associations with abstract concepts can inform design decisions in both hard- and software user interface design.

Kurzfassung

Der Einsatz von Farbe im User Interface Design ist sowohl Kunst als auch Wissenschaft. Designer legen häufig ästhetische Aspekte der Farbgestaltung in den Fokus, aber vernachlässigen dabei, dass Farbe ein wichtiger Bedeutungsträger ist, der tiefgreifende psychologische Konsequenzen mit sich bringt. Die Farbpsychologie versucht diese Lücke zu füllen, steckt jedoch noch in den Kinderschuhen. Vor allem mangelt es an einem theoretischen Ansatz, der semantische Farbassoziationen vorhersagt und erklärt, die von einer Vielzahl von Menschen in einer Vielzahl von Kontexten geteilt werden.

Vor diesem Hintergrund wird in dieser Arbeit die *Conceptual Metaphor Theory of Color* (CMToC) entwickelt, welche Vorhersagen über kulturübergreifende und erfahrungsbasierte semantische Farbassoziationen trifft und diese unter Berücksichtigung von empirischen Daten erklärt. Die Theorie beruht auf der Idee aus der kognitiven Linguistik, dass die metaphorische Sprache wertvolle Einblicke in unsere mentalen Modelle in Bezug auf Farbe liefert. Aus der Diskussion dreier Arten von konzeptuellen Metaphern, die Farbassoziationen mit physischen und abstrakten Konzepten abdecken, werden schließlich empirische Forschungsfragen abgeleitet.

Die erste Forschungsfrage befasst sich mit der Verwendung von Farbe zur Vermittlung von physischen Informationen wie *Gewicht* im User Interface. Die Ergebnisse von vier Online-Befragungen mit insgesamt 295 deutschen und japanischen Teilnehmern ermitteln den Einfluss von Farbton, Sättigung und Helligkeit auf Assoziationen mit 16 physischen Eigenschaften. Zwei Drittel dieser Farbassoziationen werden von der CMToC korrekt vorhergesagt. Die Teilnehmer aus beiden Kulturen unterschieden sich dabei nicht wesentlich in ihrer Zuordnung von Farben zu physischen Eigenschaften.

Die zweite Forschungsfrage befasst sich mit der Verwendung von Farbe zur Vermittlung von abstrakten Informationen wie *Wichtigkeit* im User Interface. Eine experimentelle Studie mit insgesamt 75 deutschen und japanischen Teilnehmern validiert hierzu Assoziationen zwischen Farben und abstrakten Konzepten in Form von Farb-Populationsstereotypen wie *wichtig ist dunkel*. Die Mehrheit dieser Farbassoziationen (86%) wird von der CMToC korrekt vorhergesagt. Auch hier unterscheiden sich die Teilnehmer aus beiden Kulturen nicht wesentlich in ihrer Zuordnung von Farben zu abstrakten Konzepten.

Die dritte Forschungsfrage befasst sich mit der Automatizität der vorhergesagten Farbassoziationen mit physischen und abstrakten Informationen als eine Voraussetzung für intuitive Benutzung. Die Ergebnisse von drei Studien mit insgesamt 85

deutschen und japanischen Teilnehmern zeigen am Beispiel *Temperatur*, dass Farbe automatisch die Identifikationsgeschwindigkeit verwandter physischer Eigenschaften beeinflusst, aber nicht umgekehrt. Farbe und abstrakte Informationen werden nicht automatisch assoziiert.

Als Ergebnis dieser Studien lässt sich feststellen, dass die Vorhersagen von CMToC kulturübergreifend gültig sind. Abgeleitete implizite Farbassoziationen mit physischen Eigenschaften und explizite Farbassoziationen mit abstrakten Konzepten erlauben es, empirisch fundierte Designentscheidungen zu treffen, die bei der Gestaltung von Hard- und Software nutzbringend eingesetzt werden können.

Acknowledgments

This dissertation was conducted at the Chair for Psychological Ergonomics at Julis-Maximilians-Universität Würzburg. I take this opportunity to express gratitude to all of those who directly or indirectly have lent me their hand in this venture.

I would like to express my boundless gratitude to Jörn Hurtienne, my advisor and mentor, for guiding me throughout this academic journey and for sharing with me his deep interest in Conceptual Metaphor Theory. His unlimited willingness to invest himself into the guidance and support of young scientists and his ability to stimulate thinking and to encourage love for your work are unrivaled. I am extremely thankful and indebted to Takashi Toriizuka for sharing his expertise, and sincere and valuable guidance and encouragement extended to me, as well as for providing me with all the necessary facilities for the research I conducted in Japan. The time he spent with me made me perceive each and every step I took in writing my thesis as an enjoyable and rewarding experience. I am also grateful to the Japanese Society for the Promotion of Science who financed my time as a visiting researcher in Japan through a pre-doctoral stipend.

The studies presented here owe much to the help of my students and colleagues. Lennart Arlt, Sina Lory, Konstantin Schall, Frederik Rössler, Katja Hünig, Miriam Hosak, Florian Haselbach and Mio Ikeda helped with gathering the data in the surveys and experiments. I especially thank Philipp Schaper for our wonderful office talks, and Frank Seyfarth for his enduring technical support. For enlightening discussions of Image Schemas I would like to thank Prof. Dr. Beate Hampe. I am grateful to Prof. Dr. Matthias Rötting, Dr. Antje Venjakob, Dr. Brett Molesworth, Nora Gosch, and Dr. Tobias Grundgeiger, who have provided valuable suggestions and comments on early and late versions of my work.

I am grateful to Wolfgang Paier for his support, encouragement and humor, and for all those days and nights of listening to my confusing ideas (he probably learned more about colors and metaphors than he would have liked). I want to thank Robert Tscharn for his loving support during the final stages of writing up, including all the cooked cheer-up meals, the morning discussions about the role of V4 in the mental representation of conceptual metaphors, the long and short writing retreats and the frequent reminders that my daily writing duties are not yet fulfilled. I am very happy to have Robert by my side. Finally, I take this opportunity to express profound gratitude from my deep heart to my parents for their continuous support throughout my life.

Contents

List of Abbrevations

ANOVA	ANalysis Of VAriance (statistical procedure)
CI	Confidence Interval (statistical parameter)
CMT	Conceptual Metaphor Theory
d	Symbol for effect size (statistical parameter)
df	degrees of freedom (statistical parameter)
EPA	Evaluation Potency Activity (Osgood, 1952)
GUI	Graphical User Interface
IT	Information Technology
M	Mean (statistical parameter)
n.d.	nicht definiert
p	probability (measure of statistical significance)
R^2	coefficient of determination (statistical parameter)
SD	Standard Deviation (statistical parameter)
TUI	Tangible User Interface
UI	User Interface

Designers, who are now more than anyone responsible for colouring our world, have a choice before them. They can continue to devalue colour by using it in an arbitrary, non-natural way, or they can recognize and build on humans' biological predisposition to treat colour as a signal. If they choose the latter, bolder course they might do well to study how colour is used in nature. Nature has, after all, been in the business of design for over a hundred million years.

Nicholas Humphrey, 1976

1. Introduction

Color is a ubiquitous perceptual stimulus that carries meaning and can impact the way we feel, think and act (Elliot & Maier, 2014). Color information is automatically processed to make inferences about the physical, and, increasingly, the digital environment. Therefore, color is an important dimension to be considered in the design of interfaces between humans and machines, as every pixel and piece of hardware seen by the user has a specific color and can impact the interaction. When applied appropriately, color can guide our attention to task-relevant parts of user interfaces (UIs), ease intuitive interaction with technology, and disambiguate the identification of physical properties as well as abstract concepts. When used inappropriately, color may impair the overall user experience. In the field of human-computer interaction (HCI), design recommendations involving color are usually restricted to issues of legibility, aesthetics or color preferences (MacDonald, 1999). But as demonstrated by recent empirical advances in the field of color psychology, there is much more to color than making interfaces more appealing and improving foreground-background contrast ratios. For example, it has been shown that the seemingly task-unrelated background color of a graphical user interface (GUI) influences the way users evaluate its affective content. Participants who judged the perceived sentiment of chat messages against a monochrome black background vs. a white interface rated the chat sentiment significantly more negative (Giron, 2016). These findings demonstrate that task-irrelevant peripheral color information can bias users' affective judgements in a way ergonomic guidelines and aesthetics do not predict. Given the widespread use of dark themed UIs, night-modes and battery saving states of technical devices, this bias can unfold its impact on an immense scale, considering the global user base.

Today, designers do not fully harness this potential of color as design element to enhance intuitive interaction with technology. In this work, intuitive use is understood as subconscious application of prior knowledge (Naumann et al., 2007). As an example, consider the context of online banking. Managing money and handling money-related sensitive data is a very private issue. Especially via the internet, doubts over the safety of this process are widespread. Thus, security-related concepts are automatically activated in the user's mind. Now, if the automatically processed color design of the online banking site fails to convey concepts like "protected" or "secure", it can negatively affect website trust, credibility, and even marketing of its business (Lindgaard, Dudek, Sen, Sumegi, & Noonan, 2011; Pengnate & Antonenko, 2013; Jahanian, Vishwanathan, & Allebach, 2015). Moreover, as color associations are considered a cultural thing, users with different cultural backgrounds might have different expectations and preferences regarding color design. But how can designers determine which colors go best with the abstract concept of security, even for an international audience?

Currently, apart from relying on their own intuition, designers can draw on "common knowledge" about color meaning, often stemming from non-scientific sources (Elliot & Maier, 2007), or following best practices (Jahanian et al., 2015). As guidelines and industry standards (e.g., indicating *cold* and *warm* water supply on a tap interface by using *blue* and *red*, respectively) are not available for many contexts, arbitrary color choices that can hamper intuitive use are the result. This unfavorable situation is forced, in part, by the lack of theoretical work in color psychology that precisely explains and predicts semantic color associations in a large variety of contexts (Sethi, Coman, & Stan, 2001; Elliot, 2015). Rather, the popular and applied literatures are replete with lay statements about symbolic color associations and their implications (Elliot & Maier, 2007), like *green stands for greed* and *red promotes determination.* Such statements almost exclusively focus on the color property hue, whereas the potential of saturation and brightness as possible design factors is left unused (Elliot & Maier, 2014; Elliot, 2015). Another drawback is that the statement itself does not tell *why* certain color properties are associated with specific concepts, limiting transferability to related but different contexts, concepts and color samples. A third shortcoming is that cross-cultural issues are often neglected, although color design plays a crucial role for adapting hard- and software to cultural requirements of different target markets (Collins, 2001). What is needed is theoretical guidance for identifying culture-independent color meanings in which color is not only a symbol with an arbitrary and conventional relation to the object it stands for (Caivano, 1998). Rather, the potential for the design of intuitive use of technology lies in experience-based color associations that are frequently encountered in our daily lives and are thus automatically activated. In such color associations, a relationship of similarity or physical contiguity exists between color as representamen and the object it indicates (Caivano, 1998; Ou et al., 2012). Thus, the goal of this work is to provide theoretical guidance for identifying cross-cultural and experience-based semantic color associations that facilitate intuitive use with technology. To this extend, image schema theory and Conceptual Metaphor Theory (CMT) within the framework of embodied cognition are applied to the field of color psychology.

In the past decade, embodied cognition has been gathering impact as a unifying perspective for psychology. Central to embodied cognition perspectives is the significant role of the body for cognitive processing and development (Brooks & Brooks, 1991; Varela, Thompson, & Rosch, 1992; Clark & Chalmers, 2010; O'Regan & Noe, 2001; Ehrlich, Levine, & Goldin-Meadow, 2006; Barsalou, 2008; Clark, 2008; Shapiro, 2009; Glenberg, 2010; Boroditsky, Fuhrman, & McCormick, 2011). Current approaches to embodied cognition can be divided into more or less radical streams in terms of which role the body plays for cognition and especially mental representation, e.g. (Shapiro, 2009). CMT falls into the stream of grounded cognition (A. Wilson & Golonka, 2013). Characteristically for this stream is that mental representations are shaped by our bodies, but not, as in more radical approaches, replaced by bodily activity. According to CMT, the physical experiences we make by moving through the material world (e.g., perceiving color) scaffold the development of our conceptual knowledge (Lakoff & Johnson, 1980). Language provides a window into this conceptual system as metaphors in speech can indicate how we think about the concepts linked in linguistic expressions. Applied to color psychology, this means that investigating the way people talk about color using linguistic metaphors can shed light on how we subconsciously relate color to other concepts in the underlying conceptual metaphors. Basically, a metaphor is a figure of speech in which one expression is used to designate something it does not literally denote. Metaphors can be described in the form of TARGET DOMAIN IS SOURCE DOMAIN. The *source domain* is the conceptual domain from which the metaphorical expression is drawn. The *target domain* is the conceptual domain that takes its structure from the *source domain* and is understood through the metaphorical link. Phrases such as *heavy red* or *I feel blue* are expressions of the metaphors RED IS HEAVY and SAD IS BLUE. These metaphors possibly reflect how color is subconsciously linked to physical properties (weight in this example) and abstract domains (sadness).

In this work, different types of conceptual metaphors are examined that are potentially fruitful for the aim of this work: to derive cross-cultural semantic color associations that are rooted in frequent experiential correspondence between color, the environment and bodily states, see Fig. 1.1 and Tab. 1.1.

Table 1.1.: *Source and Target Domains of Different Types of Conceptual Metaphors Related to Color, With Examples*

Source Domain	Target Domain	Conceptual Metaphor	Example
Color	Abstract Domain	(1) Color Metaphor	SAD IS BLUE
Image Schema	Color	(2) Image Schema-Color Metaphor	RED IS WARM
Image Schema	Abstract Domain	(3) Image-Schematic Metaphor	INTIMATE IS WARM
Image Schema Substituted by Correlating Color	Abstract Domain	(4) Color-Substituted Image-Schematic Metaphor	INTIMATE IS RED

(1) The first type concerns a mapping from the physical *source domain* color to another, more abstract, *target domain,* e.g. SAD IS BLUE, as expressed in *I feel blue.* Such *color metaphors* can leverage the design of complex ideas by physical means of color. The number of such color-to-abstract mappings that can be identified through linguistic analyses is extremely limited and therefore cannot account for the vast

amount of abstract concepts that have to be conveyed in UIs. Thus, this type of metaphor will play a minor role in this work.

(2) The second type of conceptual metaphor involves color associations with so-called image schemas. Image schemas are non-linguistic patterns of bodily interactions with the external world like WARM-COLD, HEAVY-LIGHT, or BIG-SMALL (Lakoff, 1990). Some image schemas naturally co-vary with color. For example, the color *red* often appears on heat sources like fire, the sun or blood (Ho, Iwai, Yoshikawa, Watanabe, & Nishida, 2014), and is therefore associated with the image schema WARM. *Image schema-color metaphors* like RED IS WARM as expressed in *warm red* link a physical *source domain* (weight) to the physical *target domain* of color. Such associations between colors and image schemas can be used to guide design decisions concerning physical concepts. For example, warm temperatures match best with and can even be substituted by long wavelength hues like red.

(3) The third type of conceptual metaphor that might prove to be valuable in predicting color associations with abstract concepts are image-schematic metaphors like INTIMATE IS WARM. This is because the human mind re-uses image schemas for structuring abstract concepts (Johnson, 1987; Lakoff & Johnson, 1980, 1999). Image-schematic metaphors are plentiful, bridge the gap between physical properties and abstract concepts and so far had been successfully applied in user interface design (Hurtienne & Blessing, 2007). In these metaphors, the *source domain* can be instantiated through color, e.g. INTIMATE IS RED. Such (4) *color-substituted image-schematic metaphors* cannot be directly accessed through language as they are obtained from image-schematic metaphors like INTIMATE IS WARM that do not have color in *source* or *target domain*. Rather, the image schema in the physical-to-abstract mapping is replaced by a color it naturally correlates with, e.g. warm - red.

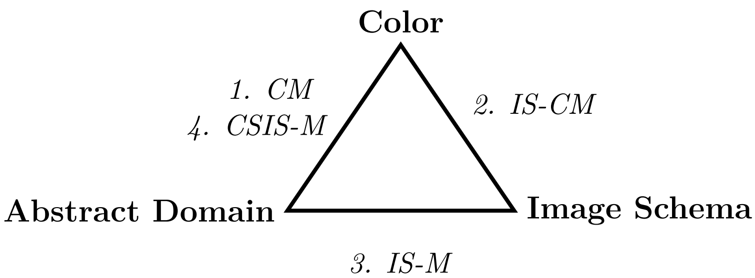

Figure 1.1.: Different types of conceptual metaphors that are fruitful to derive cross-cultural semantic color associations: 1. CM (Color Metaphor), 2. IS-CM (Image Schema-Color Metaphor), 3. IS-M (Image-Schematic Metaphor), 4. CSIS-M (Color Substituted Image-Schematic Metaphor).

The mappings of the last type do not only cover a broad variety of contexts for GUIs, but also bring advantages for the design of tangible user interfaces (TUIs). TUIs are graspable interfaces that enable the user to interact with digital information through manipulating the physical environment (Ishii, 2008). Recent trends like shape-changing interfaces and actuated tangibles unfold new possibilities to interact with technology by making use of the human ability to manipulate physical objects. However, dynamic changes in physical properties like size, temperature or weight are oftentimes expensive and impractical, so that designers are looking for replacement options (Antle, 2011). Substituting haptic instantiations of image schemas by color in image-schematic metaphors could be a practical alternative, as color is very easy to manipulate through displays and projections on(to) tangible objects (Löffler, Arlt, Toriizuka, Tscharn, & Hurtienne, 2016).

Another general advantage of basing color design on three introduced types of conceptual metaphors draws from the fact that these metaphors represent a form on *sensorimotor* prior knowledge. Possible sources of prior knowledge have been categorized by Hurtienne and Blessing (2007) into four levels along a continuum from innate, to sensorimotor, cultural and expertise knowledge. The extent to which a group of people share a color association is greatly dependent on the source of prior knowledges it stems from. Color associations that are supposed to originate from the lower levels in the continuum of prior knowledge, mainly *innate* or *sensorimotor*, should be available to a broad range of people, e.g., with different cultural backgrounds (Hurtienne & Langdon, 2009). Furthermore, they should be automatically triggered because knowledge at these levels is so fundamental and frequently encoded as well as retrieved in a large variety of situations. Consequently, basing design decisions on pervasive color associations can foster the development of a universal and robust color-meaning vocabulary that is intuitively to understand by the majority of users. In this work, the promises of such a Conceptual Metaphor Theory of Color (CMToC) to predict experience-based and culture-independent color associations for a variety of contexts are theoretically discussed and empirically investigated. The scope of this work and its structure are introduced in the following subchapters.

1.1. Scope

The goal of this work is to develop a CMToC that

- predicts concrete cross-cultural semantic associations of color with physical properties and abstract concepts to accommodate for different needs in hard- and software UI design and to inform design decisions,
- explains why specific semantic color associations occur based on empirical data to facilitate the applicability in a variety of contexts
- facilitates the design of intuitive use through semantic color associations that are supposed to be processed automatically and subconsciously.

The following topics need to be addressed beforehand:

- Current theoretical considerations in color psychology need to be reviewed, compared in terms of their predictive and explanatory power regarding semantic color associations, and analyzed for weaknesses.
- Premises of CMToC need to be developed by applying image-schema theory and CMT to color psychology.
- Existing empirical evidence for metaphors involving color needs to be reviewed and limitations of these studies need to be pointed out.

After the theoretical basis is set, empirical studies are conducted that address whether the semantic color associations with physical properties and abstract concepts predicted by CMToC hold across cultures, can be applied in different contexts and are automatically activated. Once these goals are met, the foundation is laid to derive practical recommendations for color design in HCI.

1.2. Chapter Overview

Chapter 2 reviews current theoretical work in color psychology. Starting from the perspective of evolutionary psychology, the importance of color as perceptual stimulus that carries meaningful information to the observer is worked out. Then, three theoretical approaches are compared regarding their power to predict and explain semantic color associations that can support the design for intuitive interaction. The chapter concludes with identifying the need for more explanatory rather than descriptive work in the area of color psychology to facilitate its applicability in HCI.

Chapter 3 introduces CMT and image schema theory and applies these to the area of color psychology. Three types of conceptual metaphors that potentially predict color associations are discussed: color metaphors, image schema-color metaphors, and color-substituted image-schematic metaphors with a focus on the latter two as they cover a broad range of color associations. To assess the cultural stability of predicted color associations, their sources of origin are classified along a continuum of prior knowledge.

Chapter 4 develops CMToC based on the idea that frequently encountered experiences and bodily constraints form mental models involving color, which are sometimes surfaced in metaphorical language. Existing empirical evidence on color metaphors, image schema-color metaphors, and color substituted image-schematic metaphors is reviewed and limitations of previous studies are discussed. Finally, it determines the scope of the empirical part of this work.

Chapter 5 reports the first set of empirical studies. It presents four cross-cultural surveys that investigate 16 image schema-color metaphors. Study 1 aims at quantifying the image schema-color relationships in German subjects. Resulting regression models predict the relative impact of each color attribute on the perceived image-schematic dimension. Study 2 is a direct replication of the first survey with Japanese participants to access the cross-cultural stability of image schema-color associations. The following two surveys apply image schema-color metaphors to two-color samples and determine color population stereotypes with German (Study 3) and Japanese (Study 4) subjects. In a review of all four studies it is discussed to which extend image schema-color metaphors are culture-independent and easy to transfer to different contexts.

Chapter 6 experimentally evaluates selected color substituted image-schematic metaphors. In Study 5, metaphorical extensions of BIG-SMALL, HEAVY-LIGHT, STRONG-WEAK and WARM-COLD are tested with German and Japanese subjects in form of color population stereotypes. The chapter closes with a discussion if haptic-color substitution is a valid approach for the design of TUIs and to which extend color-substituted image-schematic metaphors are culture-independent.

Chapter 7 presents three experimental studies that test whether image schema-color metaphors and color-substituted image-schematic metaphors are processed automatically on the example of the WARM-COLD image schema and related metaphorical extensions. This processing characteristic is a prerequisite for achieving intuitive use. In Studies 6 (Japanese subjects) and 7 (German subjects) investigate whether processing image schemas and abstract domains automatically influences the identification speed of color in a Stroop paradigm. Study 8 (German subjects) utilizes a reverse Stroop task to test whether perceiving color automatically influences the identification speed of related image schemas and abstract domains. In a summary the practical implications of these findings are discussed.

Chapter 8 wraps up the findings of the theoretical and empirical studies presented. The chapter summarizes and evaluates the contribution CMToC makes to predict semantic color associations to facilitate the design for intuitive use compared to other theoretical works in color psychology. The chapter closes with pointing to open questions for future research.

The Appendix contains material from the experimental studies and detailed results of the data analyses.

2. Current Work in Color Psychology

2.1. Multidimensionality and Meaningfulness of the Color Signal

Color is one of the most fundamental variables in vision. However, color is not a physical object property, but the visual signal of light reflected by surfaces and interpreted by our brains (Newton, 1730). Caivano 1998 sums up color perception in a nutshell (p. 396):

> *It is a well-known fact that color is not in the physical matter, nor in the luminous radiation; it is an image now we can say a sign produced in the mind of an organism equipped with a sensory system, vision, that reacts to a certain portion of that radiation. This image or sign is the reproduction that the visual system makes of the radiation coming from the light sources or from the objects reflecting or transmitting that radiation. This is the most primary function as a sign that color performs, that is to say, that function by which color works as a substitute for the physical radiation in order to carry to the brain useful information about the external world.*

The color signal varies among the basic dimensions of hue, saturation and brightness (Fairchild & Fairchild, 2005). This multidimensionality of color is captured by so-called *color models*. The most common color models are divided into the categories *perceptually uniform*, *device-derived* and *intuition-based* (Rheingans, 1997). Examples of perceptually uniform color models include CIELAB and Munsell. Differences of color parameters in these color spaces correspond to differences in the perceptual distance between two colors. In CIELAB, the color components represent the physiological structure of opponent-color channels in the visual pathway. In contrast to perceptually uniform color spaces, the color components in video device-derived color models relate to the signals of display devices. A widespread example of such a model is RGB (Red, Green, Blue), which is used for computer displays. According to Rheingans, the advantage of defining color on video device-derived models is that no additional transformations need to be conducted in order to display them on computer screens. Finally, intuition-based color models aim at providing a more intuitive way for people to specify colors by defining hue as basic color quality. Hue is the color property after which colors are named and is central in differentiating between colors. The place of a color in an intuition-based color model is assigned based on hue. A common example of an intuition-based color model is Hue-Saturation-Brightness (HSB). HSB defines hue values in the range of 0-360 degree as angular distance from red (Rheingans, 1997). The vividness of color is considered in the saturation component along the x-axis of the color space from 0 to 100%, and the third color dimension defines the amount of emitted light (brightness) along the Y-axis of the color space from 0 to 100%. The major drawback of both device-derived and intuition-based color models is that the euclidean distance between colors in the color space is not proportional to the perceived difference. However, even for perceptually uniform color spaces, when comparing colors that lie across a large distance in color space, the linear relationship between geometric distance and perceived difference is less accurate. For this reason, the HSB color space is the main color model used in this work, because it closely matches our conceptual structure of color and the values are easiest to interpret for the HCI professional. Nevertheless, the HSB values can be transformed into for example CIELAB values that capture the perceptual structure of human color vision better or RGB values for an application involving display devices simply by using an online color converter.

From the perspective of evolutionary psychology, color vision was naturally selected for survival because the color signal carries important information for the observer. Although the evolutionary advantage of color vision is still under debate (Kremers, Silveira L C L, Yamada, & Lee, 1999; Gegenfurtner & Rieger, 2000), it is assumed that color plays a role in discriminating figure from ground, depth perception and object recognition (Marr, 1982; Legge & Chung, 2010). In scene recognition and retrieval, colors contribute to fast scene recognition when they are predictive, or diagnostic, of the pictured scene category. For example, the recognition of a beach scene is facilitated when the scene is appropriately colored, i.e. yellow sand, blue sea and green palm trees (Oliva & Schyns, 2000). This effect can not be observed for color-non-diagnostic scenes (e.g. an urban landscape), luminance-controlled gray-scale images and for inappropriately colored scenes (a beach scene with blue sand, brown sea and red palm trees) (Gegenfurtner & Rieger, 2000). This selective advantage of color for scene and object recognition can be explained by the identity-diagnostic quality of color (Cant, Large, McCall, & Goodale, 2008). In nature, color often indicates certain physical properties as well as the chemical composition of materials (Caivano, 1998). Moreover, object memories are often defined by color (Lewis, Pearson, & Khuu, 2013) as an image characteristic that enhances recognition memory (Onasanya, 2002; Garber, Hyatt, & Boya, 2008). Thus, object recognition and retrieval are facilitated if color serves as a reliable cue for object identity and state (Tanaka & Presnell, 1999), e.g. edibility (Wheatley, 1973; Clydesdale, 1993). However, this does not apply to arbitrary colored man-made objects (Humphrey, 1976) normally encountered in urban scenes, where color is not indicative of material and surface properties. This means that man-made objects like hard- and software UIs can take advantage of the facilitating effects of color for recognition and retrieval if they make use of the identity-diagnostic quality of color based on stable statistical structures of the natural environment.

Aside from the advantage of the color signal for general object recognition and retrieval, trichromatic color vision is considered to have been evolved for specialized uses that are relevant for humans (Humphrey, 1976). Examples are the detection of fruit (Gegenfurtner & Rieger, 2000) or recognizing differences in skin color of conspecifics accompanying emotional, sexual and threatening states (Changizi, Zhang, & Shimojo, 2006). The importance of such natural or biologically-based color associations also emerges when people are asked to name their most prominent color associations and feelings for different colors. Most of the answers draw parallels to nature (blue of the ocean, yellow of the sun, etc.) (Naz & Epps, 2004), as the statistical occurrence of color in nature is much more reliable than in man-made environments. Moreover, preference judgements for real-world images are also found to be more robust and consistent across subjects compared to abstract images (Vessel & Rubin, 2010). This is not surprising considering that the mammalian visual systems' anatomy matches the statistics of the natural environment (Field, 1987). Natural scenes typically have low chromatic contrasts (Burton & Moorhead, 1987; Howard & Burnidge, 1994) and in order to detect subtle nuances, humans are highly sensitive to hue, consistent with a high prevalence and high color contrast-gain of retinal cells (Chaparro, Stromeyer, Huang, Kronauer, & Eskew Jr, 1993). Chaparro et al. therefore conclude that *"colour is what the eye sees best"*. Therefore, the visual system evolved to see color and simultaneously constrains which colors can be seen through its unique anatomy (Varela et al., 1992).

Although color is an important carrier of information in natural environments, the modern human is also frequently exposed to artificially colored objects and scenes. Kaufmann studied the prevalence of typical usages of the word *red* in a corpus-based study with over 6000 excerpts from newspaper articles on politics, culture, sports etc. (Kaufmann, 2006). Only one third referred to naturally colored objects, whereas two thirds referred to artificially colored media media. Humphrey describes this circumstance as *debasement of the color currency of nature* (Humphrey, 1976). He claims that humans always seek meaning in *nature's color-coded messages* as a consequence of their evolutionary heritage, even if no meaning is intended. Thus, even if color is used indiscriminately on man-made objects, it is interpreted as a meaningful signal based on biological predispositions, influencing perception, cognition and behavior (Elliot, Maier, Moller, Friedman, & Meinhardt, 2007; Gil & Bigot, 2015). Over decades empirical findings on the psychological effects of colors accumulated, far too many to cover them in a single review. Findings span, among many others, diverse areas like gustatory judgements of the taste of wine influenced by ambient light color (Oberfeld, Hecht, Allendorf, & Wickelmaier, 2009), perceived download speed of websites influenced by background screen color (Gorn, Chattopadhyay, Sengupta, & Tripathi, 2004) and effects of indoor lighting on the performance of cognitive tasks and interpersonal behaviors (Baron, Rea, & Daniels, 1992). It is not surprising that researchers also investigated if colors, in the absence of contextual cues, evoke some default physiological effects. For example, Goldstein (Goldstein, 1942) proposed color effects as a function of different wavelengths, with long wavelength colors (e.g. red) exhibiting an activating effect and short wavelength colors (e.g. blue) exhibiting a calming effect on organisms. However, empirical support for context-independent effects of color perception on physiological parameters is mixed (Kaiser, 1984; Jacobs & Hustmyer Jr, 1974; Caldwell & Jones, 1985; N. J. Stone & English, 1998; Fehrman & Fehrman, 2000; A. Wright, 2001).

To summarize, it has been illustrated that color perception has likely evolved because scene and object recognition as well as memory retrieval benefit from the predictive character of color regarding material properties and object identity. Thus, color is perceived and interpreted as a meaningful environmental signal, constrained by the anatomy of human color vision, and has plentiful physiological and psychological effects. The basic properties of color hue, saturation, and brightness are captured in color models. Hue is the most-studied color characteristic. This is not surprising, as hue directly translates as wavelength, which is the most psychologically recognizable color characteristic and also what most people think of when hearing “color”. Thus, colors are, in the first place, named after hues. However, brightness and saturation play also a major role in object recognition and have extensive psychological implications (Elliot & Maier, 2014). The following example in Fig. 2.1 should illustrate what can be communicated through the three dimensions of color.

In the upper photo, the luminance channel (brightness) is turned off, making the content of the picture barely recognizable. In contrast, the middle photo only contains information from from the luminance channel, as hue and saturation are turned off. As a result, the content can be recognized as sushi on a plate, despite the lack of hue and saturation. By looking closely, one notices that the lower piece of sushi looks slightly darker. But only in the fully colored photo at the bottom, the lower piece of sushi can be identified as being less fresh than the upper piece, as the freshness of food is often indicated by both high brightness and high saturation. Thus, in this example, brightness is the main facilitator in object recognition and the identification of object properties (freshness of fish), while hue and saturation support object recognition through figure-background contrast. The identification of the object property ’freshness of fish’ is further disambiguated through saturation information, increasing the likelihood that the upper piece is preferred. This example shows that it is important to study the effects of color on affect, cognition and behavior as a whole - and color psychology is the field that endeavors to do so.

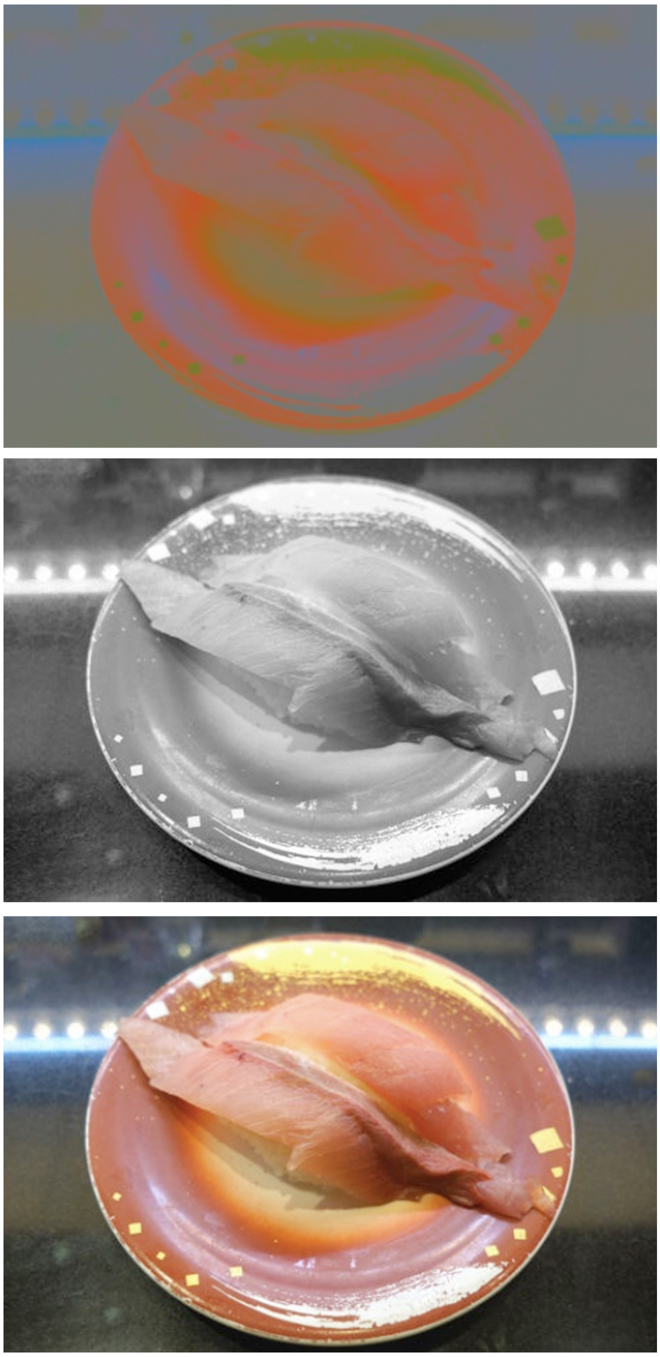

Figure 2.1.: Top: Photo without luminance information. Middle: Photo without information of hue and saturation. Bottom: Photo with hue, saturation and luminance information.

2.2. Theoretical Frameworks in Color Psychology

In their recent *Annual Review of Psychology*, Elliot and Maier state that color is an extensively studied phenomenon in terms of color physics, color physiology, color linguistics and categorization, color appearance phenomena, color deficiency and color reproduction (Elliot & Maier, 2014). However, theoretical work that accounts for how and why color affects cognition and behavior is scarce, as pointed out by several scholars (Levy, 1984; Whitfield & Whiltshire, 1990; Fehrman & Fehrman, 2000; Elliot et al., 2007; Elliot & Maier, 2014). Elliot et al. 2007, p. 13, put it in their article "Color and psychological functioning: the effects of red on performance attainment":

> *Individuals encounter a kaleidoscope of color in navigating daily life. Surprisingly, almost nothing is known at present regarding how the different colors that people perceive impact their affect, cognition, and behavior.*

Research in the area of color psychology slowly accumulated over the past decade and the following section reviews and compares the three main theoretical advances that have potential to predict a broad range of semantic color associations. Broad focus was preferred over collecting very specific conceptual statements about color effects in certain contexts. Findings such as that the perception of red signals dominance and enhances performance in sports (Hill & Barton, 2005) offer, on the one hand, precise conceptual propositions in a defined but narrow context. However, it is not clear how this specific hypothesis can be connected to a more HCI-related application. Moreover, having a plethora of individual findings that account for only very specific conditions is rather impractical compared to revealing basic mechanisms from which hypotheses for a variety of contexts can be deducted. In the following, the three selected theoretical works are reviewed with a particular focus on their predictive and explanatory power regarding semantic color associations.

2.2.1. Color-In-Context Theory

A recent theoretical account of the effects of color on cognition and behavior is Elliot et al.'s color-in-context theory: a general model on color and psychological functioning (Elliot et al., 2007; Elliot & Maier, 2012a; Elliot, 2015). The authors consider both biologically based and learned sources of color meanings and effects and stress the importance of the contextual embedding of color. The six core premises of the theoretical framework are that:

- colors can carry specific meanings and are not only about aesthetical values
- color meanings are grounded in two basic sources: biologically based response tendencies to particular colors in particular situations and learned associations

that arise from a life-time of frequently encountered pairings of colors and specific semantic information

- perceiving color evokes evaluative processes of the hostility or hospitality of the stimuli (Elliot & Covington, 2001)
- evaluative processes produce motivated behavior: positive meaning results in approach responses, whereas negative meaning results in avoidance responses
- the influence of color on psychological functioning is typically automatic without conscious awareness
- color meanings and effects are context-dependent

The authors empirically tested hypotheses derived from the premises of the color-in-context theory. For example, they examined if red impairs performance in achievement contexts (Elliot et al., 2007; Elliot & Maier, 2012b). Guided by the theoretical framework, they propose that the perception of red impairs performance in achievement contexts because it is associated with danger and psychological failure. Perceiving red in achievements contexts is assumed to activate psychological responses consistent with its associations. The association between red and danger is presumed to have biologically based and learned sources: red skin is a signal for the attack readiness in conspecifics and the color of shed blood (Changizi et al., 2006). Red is also culturally associated with danger: stop signs, warning labels, traffic lights and so on. Thus, in achievement contexts, red is hypothesized to signal danger (hostility), automatically activating motivated behavior to avoid failure, which in turn undermines performance (Elliot & Maier, 2012b). The assumptions that red in achievement contexts impairs performance and that this happens in an automated fashion were empirically tested and confirmed (Elliot & Maier, 2012b).

The color-in-context theory establishes important assumptions about how color unfolds its impact on cognition and behavior. These assumptions can guide research in the field of color psychology: first, the main sources of color meanings and effects and the context in which they operate need to be identified. Although researchers can draw on some existing well-documented links between certain hues, meanings, and effects, the majority remains speculative and requires further extensive literature research and experimentation. Especially research in the most important color dimensions of brightness and saturation (Adams & Osgood, 1973; Valdez & Mehrabian, 1994; Crozier, 1997; Hupka, Zaleski, Otto, Reidl, & Tarabrina, 1997) is severely underdeveloped (Elliot & Maier, 2014; Elliot, 2015).

Color-in-context theory stresses that color effects are context-dependent. This important assumption is overlooked in most of the literature on color research, which usually aims to identify default meanings of color in absence of contextual cues, although default color meanings seem to make up, at best, a minority of color associations (Elliot, 2015). Discovering under which (psychological) context conditions color conveys meaning can illuminate the ways in which our interactions with the colored environment shape our mental lives and is a major topic for future research in the area of color psychology.

After plausible links between color-meaning and related effects have been unveiled, researchers have to work out the details how color might effect cognition and behavior in specific contexts. The color-in-context theory describes the impact of color on psychological functioning as the continual engagement in basic appraisals of stimuli as beneficial or detrimental (Elliot, Eder, & Harmon-Jones, 2013), resulting in approach-avoidance motivation. Although the authors themselves exemplify the predictive power of their theory on the well-established example of the effects of red in performance-related contexts (Elliot & Maier, 2012b), empirical results are mixed, as some scholars found the opposite: red enhanced, rather than impaired, performance (Hatta, Yoshida, Kawakami, & Okamoto, 2002; Kwallek & Lewis, 1996). These results are frequently discussed from the perspective of regulatory fit theory (Higgins et al., 1987), linking color temperature to different types of regulatory focus: warm colors are said to activate a prevention focus, cool colors activate a promotion focus, respectively (Mehta & Zhu, 2009). Thus, red does not impair performance in achievement-related contexts per se, but the task goal itself might play a crucial role (e.g. detail-oriented vs. creative).

Finally, the assumption of color-in-context theory that color effects operate automatically and usually without conscious awareness, in humans and animals alike, is well supported by literature (Meier, Robinson, & Clore, 2004; Cant et al., 2008; Zachar, Schrott, & Kabai, 2008; Takeuchi, De Valois, & Hardy, 2003; Sherman & Clore, 2009; Friedman & Förster, 2010). This processing characteristic is also an important prerequisite for achieving intuitive use.

To summarize, color-in-context theory provides a basic framework to systematically study the effects of color on cognition and behavior. However, the theory only loosely guides the identification of semantic color associations by structuring the research process from the identification of color-associations and context to the description of the automatically evoked evaluation processes that can in turn produce motivated behavior. The predictive power is thus too coarse and the approach too unspecific to derive concrete predictions for color associations with physical properties and abstract domains as needed in HCI.

2.2.2. Affective Meanings Systems

In 1952, Osgood published his work on the three fundamental dimensions of affective meaning and introduced a measuring method: the semantic differential (Osgood, 1969). The core premises of this index of meaning are described as follows (p. 227):

- “The process of description or judgement can be conceived as the allocation of a concept to an experiential continuum, definable by a pair of polar terms.” The meaning of a concept, e.g. happiness, can be quantified using bipolar rating scales of associated sensory dimensions, e.g. warm-cool. Subjects can specify the direction and intensity of their conceptual judgement by checking a position of the scale.

- "Many different experiential continua, or ways in which meanings vary, are essentially equivalent and hence may be represented by a single dimension." There are three basic underlying dimensions of meaning on which functionally equivalent and highly intercorrelated sensory dimensions vary: an evaluative dimension (good-bad), a strength dimension (strong-weak) and an activity dimension (active-passive). This is usually referred to as a pervasive semantic frame of reference.
- "A limited number of such continua can be used to define a semantic space within which the meaning of any concept can be specified." The evaluative, strength and activity dimension are sufficient to differentiate the meaning of varying concepts.

The three dimensions of affective meaning are assumed to be fundamental and pervasive factors that classify stimuli and dominate subsequent behavior: is a stimulus benevolent or threatening (evaluation E)? Is it strong or weak (strength P (potency))? Is it active or passive (activity A) (Osgood, May, & Miron, 1975)? This three-dimensional E-P-A structure has been verified and replicated in a huge number of studies, also across cultures (Heise, 1970; Osgood et al., 1975).

Although Affective Meanings Systems is not a color psychology theory per se it does predict semantic color associations. The three dimensions of affective meaning all have positive and negative endpoints. Concepts can be specified in direction and intensity of each polar endpoint of each dimension. For example, *heat* is slightly positive in the evaluation, strength and activity dimension ($E^+P^+A^+$), whereas the concept of *anger* has a similar profile but with a negative loading on the evaluative dimension, E^- P^+ A^+ (Osgood et al., 1975). The theory now predicts associations between different concepts by their corresponding polarities. Adams and Osgood applied this principle also to the color domain (Adams & Osgood, 1973). In their "cross-cultural study of affective meanings of color" they determined context-free cross-cultural E-P-A scores of the colors red, blue, yellow, green, white, gray and black as well as the concept "color" itself with 23 different samples including participants from the USA, Afghanistan, Belgium, Calcutta, Costa Rica, Delhi, Finland, France, Germany, Greece, Hong Kong, Iran, Italy, Japan, Lebanon, Mysore, Mexico, Netherlands, Sweden, Thailand, Turkey, Yucatan and former Yugoslavia. The results are depicted in Tab. 2.1.

Table 2.1.: *Cross-Cultural E-P-A Scores (Evaluation-Potency-Activity) of Different Color Terms, After Adams and Osgood, 1973*

Color Term	E	P	A
Color	+	0	+
Black	-	+	-
Blue	+	0	0
Gray	-	-	-
Green	+	0	0
Red	0	+	+
White	+	-	0
Yellow	0	-	0

Note. +: high; -: low; 0: neutral.

The E-P-A scores of the color terms were largely consistent among the investigated cultures. If at least 18 of 23 cultural probes (78%) showed the same trend, the authors interpreted the scores as universal (Adams & Osgood, 1973). They conclude that brightness (white vs. gray vs. black) and saturation (black, gray, white vs. color) are most related to changes in the affective factors evaluation, potency and activity(Adams & Osgood, 1973). The colors blue and white received the highest positive rating in the evaluation dimension, which is supported by a large number of studies. Blue is a highly preferred color and white is consistently associated with positivity (Ball, 1965; Adams & Osgood, 1973). Black and red were rated as the most potent colors while gray, white and yellow received low ratings in the strength-dimension. The authors explain this finding by the apparent weight of the colors which is highest for black and red and lowest for gray, yellow and white ((Ball, 1965; Adams & Osgood, 1973; J. Williams, Boswell, & Best, 1975)). Red is rated as the most active color, black and gray as the most passive colors (Oyama, Tanaka, & Chiba, 1962; Adams & Osgood, 1973).

Color associations are now predicted by a polarity overlap of the underlying E-P-A scores. For example, red ($E^0P^+A^+$) will be more likely associated with heat ($E^+P^+A^+$) and anger ($E^-P^+A^+$) compared to blue ($E^+P^0A^0$), because the E-P-A-scores of heat and anger are more overlapping with red than with blue. This prediction is broadly supported by literature (Tinker, 1938; Guéguen & Jacob, 2014; Balcer, 2014; Ho, Van Doorn, Kawabe, Watanabe, & Spence, 2014; Buechner, Maier, Lichtenfeld, & Elliot, 2015; Fetterman, Robinson, Gordon, & Elliot, 2011; Fetterman, Robinson, & Meier, 2012). But how to determine if specific hues are more associated with certain image schemas than others? Of course, if there is a perfect match of the related EPA-profiles, a strong association is predicted. For example, both EMPTY and gray have the identical neutral $E^0P^0A^0$ profile, thus one would expect that both concepts are associated with each other. But how to deal with EPA-profiles that only match in some parts? Logically, a half-step from "+" to "0" would be judged to be more similar than a full step to "-". But, for example, is the concept DARK ($E^-P^0A^0$) more similar to black ($E^-P^+A^-$) or yellow ($E^0P^-A^0$), as both have a match in the evaluation dimension, but a one position mismatch in potency and activity? According to early studies performed by Osgood and his colleagues Suci and Tannenbaum, the evaluation dimension always accounts for the greatest part of the explained variance, around 26% on average, whereas potency and activity account for 8% and 5%, respectively (Adams & Osgood, 1973, p. 48). Following this, black would be predicted to be more associated with DARK than yellow, because they match in the evaluative dimension, which is of greater importance.

Overall, Affective Meanings Systems allows to make a variety of predictions about color associations. By comparing the easily identifiable E-P-A profiles of colors, physical properties and more abstract concepts, hypotheses about associations can be drawn. Affective Meanings Systems explains cross-modal associations by an overlap in semantic features. However, this practical approach also has some limitations. First of all, the current 'database' of Affective Meanings Systems only includes EPA-scores for a small set of different hues, but not for saturation or brightness. The latter can only be indirectly accessed through comparing the change in EPA-scores of white, gray and black regarding brightness. While white-gray-black represent the dimension of achromatic brightness, chromatic brightness cannot be estimated. This situation is even more difficult for the effect of saturation, which can only be determined by comparing EPA-scores of chromatic to achromatic hues. However, although white, gray and black have zero saturation compared to the other hues, they seem to represent psychologically distinct concepts. Therefore, accessing the influence of saturation only through a comparison of saturated hues vs. achromatic hues can only cover a fraction of the whole image. Since saturation and brightness both play an important role in color psychology, the Affective Meanings Systems database would require an extension in these dimensions. However, and this leads to the second issue, the current available EPA-scores of color are solely based on linguistic stimuli. The color one participant imagines when reading 'red' might be different from the 'red' another participant has in mind. This would be even amplified with colors other than primary and secondary hues, such as a 'unsaturated pink'. To minimize such undesired inter-individual variation, each dimension of the EPA-score can only take one of three values: high, neutral and low. Thus, the overall measurement accuracy of EPA-scores may profit from a more fine-grained rating scale, as well as the use of less ambigue color stimuli. A third shortcoming is that how certain colors or color attributes initially acquire certain semantic features is only sparsely addressed. Adams and Osgood conclude that "[e]xplanations from all three sources - physiological, environmental, and cultural - will probably be needed to account for the apparently universal trends in affective meanings of color" (Adams & Osgood 1975, p.151). Moreover, details about the predicted color associations like the influence of context or the automaticity of the association remain unclear. Therefore, the theoretical approach cannot predict which factors influence the occurrence of color associations. The fourth limitation is mentioned by Osgood et al. 1975. Affective Meanings Systems focusses on the affective component of meaning. By using the semantic differential as method of data collection, every concept has to be rated on each scale: “TORNADO must be judged *fair* or *unfair*, MOTHER must be judged *hot* or *cold*, and SPONGE must be judged *friendly* or *unfriendly*” (Osgood, May & Miron 1975 p. 400). The words described in these examples will not be used together in their literal sense without making anomalous sentences (fair/unfair tornado, hot/cold mother, friendly/unfriendly sponge). However, participants still have to deal with these items and they do so “in very consistent ways” (Osgood, May & Miron 1975 p.400), by interpreting the terms metaphorically. In other words, the semantic differential forces metaphorical usage. As affect is dominant in metaphorical use, affective features are enhanced to the disadvantage of other features of meaning like abstractness or animacy, which can lead to an overestimation of color associations with particular concepts. Strictly speaking, Affective Meanings Systems is about affective color associations, not general semantic ones.

2.2.3. Ecological Valence Theory on Human Color Preferences

In 2009, Schloss and Palmer introduced their *Ecological Valence Theory on Human Color Preferences* (Palmer & Schloss, 2009). In this theory, color preferences are explained as the cumulative emotional response towards colored objects. For example, if bananas are one’s favorite fruit and one dislikes apples, then yellow is probably preferred over green or red. This theory is one of the first that seeks to explain why people have specific color preferences instead of only focussing on describing which colors are preferred (Ball, 1965).

The authors also report an experiment to support their claims. Participants viewed 37 different colors on a computer screen, one at a time, and were asked to brainstorm concrete objects they associate with a particular color. The participants’ answers were filtered and clustered and presented to another participant sample to rate the objects’ valence. A correlation between the weighted affective valence estimates and color preferences (for details of the analysis see Schloss & Palmer, 2010) was calculated and accounted for 67% of the variance of the data. Overall, the Ecological Valence Theory on Human Color Preferences demonstrates that not only innate and biology-based but also learned mechanisms seem to play a role in visual preference (Rentschler, Jüttner, Unzicker, & Landis, 1999). Color preferences are assumed to exhibit their influence mainly in a conscious and high-level fashion (Ho, Iwai, et al., 2014), with a remarkable similarity between human and animal color preferences (McManus, Jones, & Cottrell, 1981). Color preferences also seem to influence cognitive and behavioral aspects beyond mere product choice, such as website trust and satisfaction (Cyr, Head, & Larios, 2010). Compared to the significance of the color attributes brightness and saturation in forming peoples’ semantic color associations (Adams & Osgood, 1973; Valdez & Mehrabian, 1994; Crozier, 1997; Hupka et al., 1997), hue seems to be the most important color attribute concerning color preferences (Kereney, 1966; Palmer & Schloss, 2009).

Color preferences potentially influence semantic color associations because people are more likely to recall associations with objects that are highly liked or disliked (Palmer & Schloss, 2009), thus strengthening the association. This effect is further increased as people prefer to surround themselves with objects in colors they prefer. For example, when people are asked to rate whether a specific green is more associated with *big* or *small*, they will recall positive and negative associations with big and small green colored objects. If the most dominant or plentiful associations concern big green objects, *green* will be matched with *big*. However, even people with the same favorite color are likely to differ regarding their experiential encounters with colored objects. Therefore, no precise prediction can be deduced from this theoretical approach without having to determine associations with colored objects for each individual first, and after that, linking the objects' physical properties to colors to derive semantic color associations. In practice, such an approach is far too time-consuming and costly.

2.3. Need for a Predictive and Explanatory Framework

Many scholars demand more explanatory rather than descriptive research in the area of color psychology, emphasizing the need for a theoretical framework that does not reduce color associations to its observable manifestations, but explains and predicts how color associations affect cognition and behavior (Whitfield & Whiltshire, 1990; Bonnardel, Piolat, & Le Bigot, 2011; Elliot & Maier, 2012a; Meier, Schnall, Schwarz, & Bargh, 2012). This lack of theory may also account for the mainly inconsistent results regarding color associations and their implications (Elliot et al., 2007; Jalil, Yunus, & Said, 2012). The previous section has shown that attempts have been made to fill this theoretical gap. The three theories that were introduced seek to explain different aspects of color associations. Color-in-context theory (Elliot et al., 2007; Elliot & Maier, 2012a; Meier, D'Agostino, Elliot, Maier, & Wilkowski, 2012) is a general theoretical rationale to study the effects of color on psychological functioning. It's major limitation is that its very general predictions need to be elaborated in detail for each color association of interest, strongly limiting its applicability in HCI. Affective Meanings Systems (Osgood, 1969; Adams & Osgood, 1973) introduces the principle of polarity overlap between colors and other concepts. It predicts that associations are likely between colors and concepts with overlapping factors of affective meaning. This parsimonious explanation makes a variety of predictions about which color associations are likely to occur. However, the predictive power focusses on hue and affective color associations in general. Moreover, it lacks an explanation why colors acquire these semantic features and how they unfold their influence on cognition and behavior. Lastly, the Ecological Valence Theory on

Human Color Preferences was introduced (Palmer & Schloss, 2009). This theory explains that color preferences are likely to originate from the cumulative emotional experiences with colored objects and environments people encounter through their lifetime. As color preferences and experiential encounters with colored objects highly vary between individuals, deriving concrete predictions for user interface design from this approach is laborious and impractical. Moreover, effects of color preferences seem to operate at a conscious level, rendering them less suitable for the design for intuitive use (Meier et al., 2004; Cant et al., 2008; Zachar et al., 2008; Takeuchi et al., 2003; Sherman & Clore, 2009; Friedman & Förster, 2010).

Thus, in the following chapter, a new theoretical framework is developed that aims at predicting concrete semantic color associations and at explaining why these occur. This theoretical framework is based on work in the area cognitive linguistics: CMT and image schema theory, both under the roof of embodied cognition. Embodied cognition is currently considered to be a unifying perspective in psychology (Thomas, Joseph, & Laccetti, 2009; Glenberg, 2010) and may have potential to explain and predict the relative importance of color attributes for a wide range of color associations for both physical as well as abstract domains.

3. Conceptual Metaphors and Image Schemas

3.1. Embodied and Grounded Cognition

In recent years, embodied cognition has been gathering impact as a unifying perspective for psychology (Thomas et al., 2009; Glenberg, 2010). Embodiment describes the notion that cognition is constrained and influenced by the body and the conceptual system is grounded in observables in the world (Morrison, Oates, & King, 2001). A popular definition of the term *embodied* was introduced by Rosch et al. 1992, p.172-173:

> *By using the term embodied we mean to highlight two points: first that cognition depends upon the kinds of experience that come from having a body with various sensorimotor capacities, and second, that these individual sensorimotor capacities are themselves embedded in a more encompassing biological, psychological and cultural context.*

According to this definition, cognition is influenced by aspects of an agent's body beyond the brain itself and cannot be separated from context (also discussed under the term *situated* cognition). This embodied cognition perspective contrasts with approaches from cognitive psychology of the 1970s and 1980s and even earlier. Views as Cartesian dualism or computationalism share the notion that behavior is produced by a mind operating in terms of a closed loop system of formal symbol manipulation. One major issue with these approaches is the so-called *symbol grounding problem* (Harnad, 1990). If symbols are only defined by other symbols, it can not be explained how they acquire meaning. Embodied cognition may solve this problem by postulating that symbols acquire meaning through bodily experiences with objects and situations the symbols refer to. Cognition is not seen as a rule-based manipulation of abstract representations within a disembodied mind like in the cognitivist view (Clark, 1997). Rather, cognition it said to rely on repeated interaction with the environment in order to structure and simplify cognitive tasks (Anderson, 2003) and, in more radical streams, to "offload" cognitive work onto the environment (M. Wilson, 2002). In recent years, embodied cognition gathered central stage for theorizing and research in various areas of psychology like social psychology, emotion (Thomas et al., 2009), developmental psychology, memory, language, neuropsychology, clinical psychology and educational psychology (Glenberg, 2010), but has been rarely applied to color psychology. CMT (Lakoff & Johnson, 1980, 1999) falls into the embodied cognition sub-stream of *grounded cognition* (A. Wilson & Golonka, 2013). Grounded cognition is characterized by the importance of the body and its experiential encounters with the environment in the formation of mental representations. As CMT seeks to explain how physical as well as abstract ideas rely on bodily experience (e.g. perceiving color), it offers great potential for the prediction of semantic color associations. Because of its central role in this work, CMT is introduced in the following section.

3.2. Conceptual Metaphor Theory

In the 1980s and 90s, George Lakoff and Mark Johnson published groundbreaking work in the field of cognitive linguistics. The greatest contribution lies in stating that metaphorical linguistic expressions are only surface manifestations of underlying conceptual metaphors, which are in turn the main mechanism through which abstract concepts are comprehended (Lakoff & Johnson, 1980; Johnson, 1987; Lakoff, 1990, 1987, 1993; Lakoff & Johnson, 1999). Metaphors are defined as cross-domain mappings in the conceptual system, grounded in non-metaphorical understanding (Lakoff, 1993). For example, when people talk about *importance*, they use metaphorical expressions like "heavy matters", "he spoke weightily", "I don't attach any weight to these rumors". According to Lakoff and Johnson, the metaphorical usage of weight experiences provides insight into the existence of the underlying conceptual metaphor IMPORTANT IS HEAVY - UNIMPORTANT IS LIGHT (JÄKEL, 2003). This metaphor allows us to understand the relatively abstract and intangible concept of *importance* by importing the relational structure from the concrete experiential domain of weight (Boroditsky, 2000; Boroditsky & Prinz, 2008). As a result, the ways we reason about *importance* should show parallels with the ways we reason about weight, because weight is the source domain of the related conceptual metaphor (Casasanto, 2008).

Sensory-rich metaphors which are used to express abstract ideas are omnipresent in language and are uttered about six times every minute (Gibbs, Costa Lima, & Francozo, 2004). Asch was probably the first social psychologist who noticed about the dual nature of certain words in language. In 1955 he wrote about the physical and psychological usage of words such as *warm*, *straight* or *hard* that are used to describe both object properties as well as personality traits (Asch, 1955). Such linguistic metaphors are an important source of hypotheses about the underlying conceptual system, although some concepts are related in more complex ways than suggested by the linguistic expression (Casasanto, 2009b).

Metaphors transfer understanding from a *source domain* to a *target domain*. Because of this unidirectionality, we make use of weight adjectives to talk about importance, but do not use importance-related words to define weight-related experiences (i.e., we do not say "My shopping bags are very important [heavy]"). However, psychological studies usually find that the link between *source* and *target domain* operates bidirectionally (Meier et al., 2004; Meier, Robinson, Crawford, & Ahlvers, 2007). Having a closer look at conceptual metaphors, one might notice that the mapping is not only unidirectional, but also partly. Only certain components of the source are mapped onto the target domain. For example, while the complex metaphor CAUSES ARE TRANSFERS implies that when "she gave him a headache", she is the cause of his headache, one would not conclude that because she gave him the headache, she does not have it anymore (Lakoff, 1990; Kövecses, 2002). This partial mapping from source to target domain in conceptual metaphors is called the *invariance principle* (Lakoff, 1990). The principle describes that in a metaphorical mapping only those components of the *source domain* are mapped onto the *target domain* which remain coherent in the target context and that this set of ontological components is fixed (Lakoff, 1993).

CMT further states that some conceptual metaphors are acquired on the basis of universally observable correlations between *source* and *target* concepts and are learned obligatory during cognitive development. Such conceptual metaphors have been called *primary metaphors*, and are assumed to be universal, although culture specific exceptions exist (Lakoff, 1993; Lakoff & Johnson, 1999). Lakoff and Johnson propose that a set of embodied and prelinguistic structures central to human experience lies at the basis of many primary metaphors (Lakoff & Johnson, 1980; Johnson, 1987). These basic structures are called *image schemas*. Both image schemas and primary metaphors are introduced in the following section.

3.3. Image Schemas

Image schemas are non-linguistic, pre-conceptual patterns of bodily interactions with the external world (Johnson, 1987; Lakoff, 1987). These highly schematic and experiential gestalts emerge during sensorimotor experience while we manipulate objects or orient ourselves in the environment and integrate information from multiple modalities (Gibbs et al., 2004; Hampe, 2005a). Image schemas are even discussed to provide a basis for language acquisition by "creating an interface between the continuous processes of perception and the discrete nature of language" (Mandler, 1992), p. 587. Image schemas include spatial dimensions like FAR-NEAR and CENTER-PERIPHERY, physical force dynamic events like BLOCKAGE and BALANCE, and physical object attributes like weight (HEAVY-LIGHT), texture (SMOOTH-ROUGH) or size (BIG-SMALL), see Tab. 3.1) (Baldauf, 1997; Clausner & Croft, 1999; Hampe, 2005a; Johnson, 1987; Talmy, 2005; Lee, 2001; Hurtienne & Blessing, 2007; Hurtienne, 2011; Macaranas, Antle, & Riecke, 2012). The experiential image schemas play an important role for abstract reasoning as they motivate conceptual mappings by substituting the source domain in many *primary metaphors* (Hurtienne, 2011).

In 1997, Grady introduced the concept of *primary metaphors* as primitive metaphorical mappings with a direct experiential basis, resulting in linguistic consequences (Grady, 1997). Primary metaphors are generic cross-domain mappings in which one of

Table 3.1.: *Overview of Image Schemas Grouped by Similarity, after Hurtienne 2011*

Group	Image Schemas
Basic	OBJECT, SUBSTANCE
Space	CENTER-PERIPHERY, CONTACT, FAR-NEAR, FRONT-BACK, LEFT-RIGHT, LOCATION, PATH, ROTATION, SCALE, UP-DOWN
Containment	CONTAINER, CONTENT, FULL-EMPTY, IN-OUT, SURFACE
Multiplicity	COLLECTION, COUNT-MASS, LINKAGE, MATCHING, MERGING, PART-WHOLE, SPLITTING
Process	CYCLE, ITERATION, SUPERIMPOSITION
Force	ATTRACTION, BALANCE, BLOCKAGE, COMPULSION, COUNTERFORCE, DIVERSION, ENABLEMENT, MOMENTUM, RESISTANCE, RESTRAINT REMOVAL, SELF-MOTION
Attribute	BIG-SMALL, BRIGHT-DARK, FAST-SLOW, HARD-SOFT, HEAVY-LIGHT, SMOOTH-ROUGH, STRAIGHT, STRONG-WEAK, WARM-COLD

the two domains which are simultaneously experienced is “sensory (i.e., perceptually accessible), concrete, basic, or more clearly delineated”, the other one is “non-sensory, abstract, complex, or not clearly delineated” (Hampe, 2005b, p.44).

An often cited example is the correspondence between quantity and verticality in the primary metaphor MORE IS UP. Because we often simultaneously experience a change in verticality and a change in quantity (adding more of something usually causes the level of the substance rise), both concepts are linked to each other (Lakoff & Johnson, 1999), and this manifests itself in psychological as well as linguistic consequences. A specificity of primary metaphors is that the strength of the association is oftentimes asymmetric, that is, if the *source* domain of a primary metaphor is a gradable property, the metaphor preferably ’docks’ to one aspect of the oppositions (Grady, 2005), the so-called linguistically unmarked category. For example, we say “you are 1.74m tall” instead of “you are 1.74m short”, so, in this example, *short* is the linguistically marked category because it stands out as unusual compared to the more common *tall*. Unmarked categories are more frequent and thus receive a processing benefit (Lakens, 2012).

Grady defines three conditions for the emergence of primary metaphors: (1) primary metaphors arise from our bodily experience where we note strong correlations between (2) a sensory domain (*source*) and a non-sensory domain (*target*). (3) Both domains share a super-schematic structure, which can be for example states, scales, actions, events or trajectories. These conditions differentiate primary metaphors as a type of *correlation-based metaphors* from *resemblance metaphors*, which do not originate from simultaneous experiences, but share a super-schematic structure or feature (Grady, 1999; Ortiz, 2011).

Primary metaphors are said to be universally acquired as they arise from embodied experiences that are universal (Grady, 1997). Universal are experiences like day and night cycles, climate and seasonal changes, and tactile sensations like weight, because these physical properties of our environment did not structurally change over the course of eons (Shepard, 1984). Moreover, cultural factors and language are also important sources of metaphor acquisition (Casasanto, 2014). Because of this, some authors proposed that the repeated association between physical properties (like weight) and abstract concepts (like importance) might have caused a basic connection which is easily learned or even innate (Caporael, 1997; Damasio, 1999; A. P. Fiske, 2004; D. Cohen & Leung, 2009; IJzerman & Semin, 2010; IJzerman & Koole, 2011). Three mechanisms play a role in metaphor acquisition: evolutionary predispositions, Hebbian learning through repeated co-occurrence, or proceduralization of knowledge after extensive practice (Lakens, 2014). For example, in a recent study, Slepian and Ambady could demonstrate that 'new' conceptual metaphors can be induced in the laboratory after extensive training (Slepian & Ambady, 2014). The authors speculate that perhaps only such metaphors that have image schemas as a source domain can be learned in such a setting.

To summarize, CMT postulates that much of our thought is structured by conceptual metaphor. This basic idea is supported by an amazing amount of literature in various disciplines. Most importantly, CMT offers a unique explanatory and predictive power on perception and cognition (Landau, Keefer, & Meier, 2011) from which the field of color psychology might profit. Since many conceptual metaphors have linguistic consequences, hypotheses about links between color, other physical properties and abstract domains can be derived, of which we would else be unaware (Casasanto, 2009b; Landau et al., 2011). However, since linguistic metaphors only reflect a subset of instantiations of conceptual metaphors, empirical tests are required to reveal the exact relationships (Casasanto, 2009b). Moreover, if the origin of a linguistic metaphor can be identified among bodily, social, and cultural experiences (IJzerman & Koole, 2011), it is possible to predict its cultural stability. As conceptual metaphors are assumed to operate mostly unconscious and automatically (Lakoff, 1993), they offer great potential for the design of intuitive use. Image schemas have been shown to play an important role in conceptual thought and structure many primary metaphors. Image schemas in primary metaphors bridge the gap between physical properties and abstract concepts and therefore offer potential for hard- and software design (Hurtienne & Blessing, 2007). If some basic correspondences between color and certain image schemas can be identified, they can be used to predict associations with abstract concepts with which color does not necessarily correlate in sensorimotor experience. Thus, in the following section, premises of CMT are applied to the field of color psychology to predict cross-cultural semantic color associations.

3.4. Applying Conceptual Metaphor Theory to Color Psychology

According to CMT, metaphorical language provides hints to our underlying conceptual system. Therefore, by analyzing the way we metaphorically speak about physical properties and more abstract domains helps to uncover color associations with said concepts. However, as not all linguistic metaphors are conceptual metaphors (Casasanto, 2009b), the color associations unveiled through discourse analysis will vary regarding their impact on behavior. In order to differentiate those color metaphors that are shared by a large group of people from those which highly vary with an individual's learning history, one has to have a look at how color initially acquires meaning in different linguistic expressions related to color. Therefore, the following section categorizes the origins of conceptual metaphors along a continuum of prior knowledge and, subsequently, color associations, in order to draw conclusions about their cultural stability. The focus will be on those categories with the greatest potential to predict cross-culturally valid semantic color associations that are processed automatically, as they are the most promising for the design for intuitive use independent of the users' cultural background. Three types of metaphors are identified which stimulate hypotheses about widely shared relationships between colors and physical properties as well as abstract concepts: color metaphors, image schema-color metaphors, and color-substituted image-schematic metaphors.

3.4.1. Sources of Prior Knowledge

3.4.1.1. Origins of Conceptual Metaphors

Casasanto, 2013, distinguishes two main sources as the origin of conceptual knowledge: innateness and experience, see Tab. 3.2. Empirical evidence for innateness claims are for example provided by findings in developmental psychology. For instance, experiments revealed that some cross-domain relationships are already present at a very early age in pre-linguistic infants. For example, Dolscheid and colleagues found that four month old babies look significantly longer at cross-modality congruent stimuli (i.e., space-pitch mappings) compared to incongruent combinations (Dolscheid, Hunnius, Casasanto, & Majid, 2014). This can be interpreted as evidence that some cross-domain mappings are possibly innate. However, as Casasanto argues, "innateness claims are exceptionally hard to support experimentally. Importantly, there is no evidence that many of the mappings Lakoff and Johnson (and others) attribute to bodily experience are not innate" (Casasanto 2013, p.253).

Experiential origins of conceptual metaphors are further divided into bodily interactions, cultural and linguistic encounters (Casasanto, 2014). Universal human physiology might motivate some primary conceptual metaphors, which are only prone to little cultural influence (Kövecses, 1997; Brdar, 2001; Kövecses, 2005; McElhanon, 2006; Kövecses, 2006). An example for this is the influence of motor performance on shaping the mental representation of valence (Casasanto, 2009a, 2011; Casasanto & Chrysikou, 2011). Casasanto and colleagues performed studies in which left- and right-handers were presented with products, persons or creatures to their left and right sides and were asked to decide between the two based on different criteria. As a result, participants reliably attributed positive characteristics to their dominant side, i.e., right-handers preferred stimuli presented to their right side while left-handers preferred the stimuli on their left side. Thus, the direction of the conceptual metaphor GOOD IS RIGHT - BAD IS LEFT is influenced by how people use their hands. Another example for bodily or sensorimotor experiences as shaping factor for conceptual metaphors is the correlation of size, strength and horizontality. More voluminous agents also tend to be stronger and larger, thus the conceptual metaphors POWER IS SIZE and POWER IS UP may be at least partly determined by observed basic physical principles (Thomas et al., 2009). Such links between sensorimotor representations and conceptual processing are not static and fixed, nor do they require a lifetime of experience. Slepian and colleagues could show that even a few hours of experiencing a reliable correlation between certain concrete and abstract concepts can establish cross-domain interferences (Slepian & Ambady, 2014).

TIME IS SPACE is an example for a conceptual metaphor which concrete instantiation varies across cultures (Casasanto, 2014). Casasanto could demonstrate that subjects may spatialize time in their minds according to prevalent practices in their cultures, namely, the reading direction. Although the sagittal axis is often used when talking about temporal sequences in many languages, Dutch people are immersed in a culture in which time is linked to the lateral axis (which is common in most of the Western world). This may have lead to the development of the conceptual metaphor EARLIER IS LEFT - LATER IS RIGHT, without any linguistic manifestations. In an experimental study Casasanto and Bottini could show that the reading direction indeed influenced the mental timeline. When Dutch participants read mirror-reversed orthography, their mental timelines rotated accordingly (Casasanto & Bottini, 2014).

Finally, POLITICAL ORIENTATION IS HORIZONTALITY is an example for a conceptual metaphor that is acquired through language without any experiential correlation between political parties and spatial orientation (Casasanto, 2014). Originating from the spatial arrangement of the French Legislative Assembly, “two centuries later, liberal and conservative values are metaphorized on a left-right continuum, across many languages and cultures” (Casasanto 2013 p.12). This metaphorical mapping of the political parties’ liberality versus conservatism has been shown to have behavioral consequences. For example, in a series of reaction time experiments, Dutch participants were faster to make judgements about more liberal political parties with their left hand and faster to make judgements about more conservative parties with their right hand (van Elk, van Schie, & Bekkering, 2010). As people nowadays do not experience the historical roots of the left-right spatialization of French politicians and political depictions do not show any systematical use of left-right spatial orientation, it is likely that language created this conceptual metaphor

(Casasanto, 2014).

Table 3.2.: *Sources of Prior Knowledge of Conceptual Metaphors*

Knowledge Sources		
Casasanto 2013	Hurtienne & Blessing 2007	Spence 2015
Language	Expertise	Semantic/Linguistic Origin
Cultural Experience	Culture	Statistical Origin
Bodily Interactions	Sensorimotor	
Innate	Innate	Structural Origin

The reviewed sources of origin of conceptual metaphors (innateness, bodily and sensorimotor experiences, culture, language) are not necessarily mutually exclusive, but may reinforce each other reciprocally (Thomas et al., 2009). For example, if a conceptual metaphor is acquired through sensorimotor experience of physical laws, it can stimulate the creation of a man-made environment and linguistic expressions that are congruent with the metaphor and which then in turn contribute to metaphor acquisition. Overall, although the sources of origin might differ between conceptual metaphors, all constitute frequently repeated experience which leads to the formation of well-established prior knowledge that may be automatically retrieved. All sources discussed by Casasanto, 2013, are summarized in Tab. 3.2. Another similar classification scheme of sources of prior knowledge was put forward by Hurtienne and Blessing. They categorized sources of prior knowledge into four groups along a continuum with a varying degree of cultural dependence: innate, sensorimotor, cultural and expertise (Hurtienne & Blessing, 2007), see Tab. 3.2. For each source of origin, different assumptions hold regarding the extent to which they are present in the population. This devision of knowledge sources corresponds, in essence, to the categorization proposed by Casasanto 2013.

Innate sources refer to associations that are present "through the activation of genes or during the prenatal stage of development" (Hurtienne & Blessing 2007, p.2). This instinct-like knowledge is applied subconsciously and is universally present. Sensorimotor knowledge, the next stage of the continuum, is knowledge that is acquired from an early age on through continuous and repeated interaction with the environment. Objects and events of the environment that are based or rely on certain physical principles are rather invariable regarding the influence of cultural factors such as human interventions and creations. In addition, the basic anatomy of the human body is similar across cultures. Lakoff and Johnson distinguish these more physical and thus universal experiences that lead to the formation of image schemas from more cultural experiences, although these universal concepts are always formed in a cultural-specific environment and are more or less influenced by cultural factors (Lakoff & Johnson, 1980). Since associations of this kind stem from universal physical and social experiences, they are quasi-universal and available to a large section of the worlds' population. For example, Adams and Osgood found support for the claim that there are universal trends of attributing affect to color both in their cross-cultural study including 23 cultures as well as their literature analysis of 89 studies before 1973 (Adams & Osgood, 1973). Because associations at the sensorimotor level are very frequently encoded, they are highly automatic and, at most, subconscious. Stage three of the continuum is about cultural knowledge. Cultural-specific prior knowledge is specific to individuals of the same culture and language. Knowledge at this level can be both conscious and subconscious. Boroditsky and colleagues found that culture and language are determining factors in conceptual development. Especially abstract concepts as target domains are linked to different (but sometimes related) source domains in different languages and speakers of these languages think according to the metaphor in speech (Boroditsky et al., 2011). In line with this notion, Shankar and colleagues compared color-flavor associations between groups of participants from different cultural backgrounds. The authors found that since the participant subgroups were systematically exposed to specific color-flavor co-pairings in their cultural environment, they developed different semantic meanings (Shankar, Levitan, & Spence, 2010). For example, while 70% of the British subjects associated a dark brown drink with Cola, the post prominent association among the Taiwanese subjects was grape juice (40%). Lastly, associations at the level of expertise knowledge are the most specific and are available to an even more limited section of the population which is exposed to certain experiences or tools in their environment. Again, prior knowledge at this level can be conscious and subconscious. In the next section, the classifications of knowledge sources of conceptual knowledge and metaphors are applied to the acquisition of color associations, substantiated by existing literature.

A third classification scheme for the origin of cross-modal correspondences was proposed by Spence, 2015: the structural account, the statistical account and the semantic account. According to Spence, the structural account covers associations that originate from human body physiology and "reflect the inherent structural, wiring, or processing constraints of the human brain" (Spence et al., 2015, p.8). This approach is identical with the category *innate* put forward by Casasanto, 2013, and Hurtienne and Blessing, 2007. The statistical account covers associations that "reflect the internalization (in the observer) of the statistical regularities of the environment" (Spence et al., 2015, p.8), subsuming the sensorimotor and cultural knowledge levels distinguished by other authors. An example provided by Spence and colleagues illustrates the statistical account. If a colored cup is frequently used to serve a specific beverage, the color of the cup will be associated with the beverage it is containing - already in six month old infants (Spence et al., 2015). This acquired association is a result of the specific statistical relationship in the observers' environment, but it is not shared by many other individuals, like for example the association between *warm* and *red*. Lastly, the semantic or linguistic account explains color associations that are acquired through language and is "based on the observation that we often use the same descriptors for qualitatively different sensory impressions" (Spence et al., 2015, p.8), similar to the categories language (Casasanto, 2013) and expertise (Hurtienne & Blessing, 2007). For example, the word *warm* is used to talk about colors and temperature alike.

3.4.1.2. Origins of Color Associations

The major division of sources of prior knowledge in innate (or biology-based) and experiential (or learned) is also shared by scholarly work in color psychology (Mahnke, 1996; Aslam, 2006; Elliot et al., 2007; Sable & Akcay, 2010; Labrecque, Patrick, & Milne, 2010). One of the most well-known advocates of the biological origin of color meaning and effects is Humphrey, who has sought to trace back color effects to an innate or instinctive basis, claiming that color vision has evolved to enable humans to detect and interpret nature's color-coded messages (Humphrey, 1976). Multiple scholars discuss such an evolutionary heritage or at least an evolutionary preparedness to learn specific color meanings easily, mainly focussing on the color red (Elliot & Maier, 2014). For example, red has an appetitive character in mating contexts (Meier, D'Agostino, et al., 2012) and the association between perceptual redness and anger probably originated from the testosterone-based increase in blood flow that is most visible in facial regions (Changizi et al., 2006; Drummond, 1997). In a more recent review article, Labrecque and colleagues use the term 'embodied' to describe such wavelength-dependent biologically based color associations, claiming that they are related to arousal and exhibit their influence independent of context (Labrecque et al., 2010).

Given that it is difficult to find experimental support for innateness (Casasanto, 2014), the majority of theoretical statements on the origin of semantic color associations focusses on learning mechanisms, often defined as encounters with explicit and implicit pairings between color and semantic information and experiences throughout lifetime (Wheatley, 1973; Elliot et al., 2007; Garber et al., 2008). An important distinction regarding learned associations is made between associations that stem from 'universal human experiences' and 'culture-specific variables' like language (Adams & Osgood, 1973; Hupka et al., 1997). This distinction is similar to the category differentiation put forward by Casasanto (2013) (body, culture and language) and the sensorimotor versus cultural level of prior knowledge by Hurtienne and Blessing (2007). A thoroughly investigated example for an universal human experience as origin of a semantic color association is the link between *darkness* and negative emotions like fear as the result of our evolution as diurnal animals (Mead & Baldwin, 2001; Adams & Osgood, 1973; J. Williams et al., 1975; Shepard, 1984; Wierzbicka, 1990; Farrington & Welsh, 2008; Lakens, Semin, & Foroni, 2012; Kort, Ijsselsteijn, Haans, Lakens, & Kalinauskaite, 2010). This association is also expressed in and likely reinforced through metaphorical language like *a dark vision* or *he is in a dark mood.* On the contrary, an example for a more culture-specific color association acquired through language (Duckitt, Wall, & Pokroy, 1999) is the *red thread* as a symbol for a common theme. A speaker of the language with the *red thread* metaphor has no experiential basis of the origin of the expression anymore (*red thread* is said to originate from literary works like the Greek myth of minotaur's labyrinth or Goethe's Elective Affinities[1]. In the prior knowledge continuum proposed by Hurtienne and Blessing, the most exclusive level of prior knowledge is expertise knowledge. With regard to the origin of color associations, this level could be interpreted as language (Casasanto) or the linguistic account (Spence), or even seen more narrow, referring to color associations that are only shared by an even smaller group of people, e.g. by those who encountered specific fashion trends or brand designs (Mahnke, 1996). Tab. 3.3 summarizes the different knowledge sources of semantic color associations.

Table 3.3.: *Sources of Prior Knowledge of Color Associations, with Examples*

Knowledge Source	Explanation	Example	References
Language and other forms of expertise knowledge	Semantic color associations that originate from language or encounters with man-made artifacts, e.g. fashion, advertisement, branding.	red - structure (red thread) red - Coca Cola	(Mahnke, 1996; Hupka et al., 1997; Duckitt et al., 1999)
Cultural experience	Semantic color associations that originate from experiences shared by individuals of the same culture.	red - communism	(Adams & Osgood, 1973; Mahnke, 1996; Hupka et al., 1997)
Sensorimotor/ bodily interactions	Semantic color associations that originate from basic physical and social experiences.	red - fire	(Adams & Osgood, 1973; Wierzbicka, 1990; Mahnke, 1996; Hupka et al., 1997; Elliot et al., 2007; Sable & Akcay, 2010; Labrecque et al., 2010; Wan et al., 2014)
Innate	Semantic color associations that are inborn or easily acquired through evolutionary preparedness.	red - anger	(Humphrey, 1976; Adams & Osgood, 1973; Mahnke, 1996; Elliot et al., 2007; Labrecque et al., 2010)

Identifying the source of origin of a specific color association helps to assess its cultural stability. The deeper the color association is anchored within the knowledge

[1] https://www.redensarten-index.de, last accessed 15.03.2017

continuum, the more independent it is of the users' varying cultural background. With this regard, color associations that are innate or initially originated from sensorimotor knowledge should be rather stable across different cultures, and are therefore especially interesting for cross-cultural UI design. Moreover, color associations at these levels of prior knowledge are very frequently encoded in and retrieved from memory. Thus, they are supposed to be processed automatically, which is a prerequisite for achieving intuitive use (Hurtienne & Blessing, 2007). However, given the plethora of people's color associations, it remains a challenging research question which associations are activated in a given situation and how they affect subsequent behavior. Having identified color-associations at different levels of prior knowledge, the next section will have a look at how these color associations are expressed in metaphorical language, focussing on color associations at the lower levels of the continuum. As CMT states, conceptual metaphors are often manifested in linguistic expressions, thus can be interpreted as hints to the underlying conceptual system. Since colors are frequently co-experienced with physical properties as well as abstract concepts, it can be expected that links establish which are expressed in metaphorical language.

3.4.2. Color Metaphors

The first conceptual metaphor type of interest unveiled through the study of linguistic expressions consists of a mapping from the physical *source domain* color to another, more abstract, *target domain.* Almost all of these color metaphors concentrate on the color properties hue and brightness. Examples of color metaphors are introduced in the following.

3.4.2.1. Metaphorical Meanings of Different Hues

Metaphorical expressions involving color are ubiquitous. In idioms like *to be in the black* (being successful or profitable), *out of the blue* (to appear or happen quite suddenly), *nicht das Gelbe vom Ei* (not the yellow part of the egg, German for not the best part), *das Blaue vom Himmel versprechen* (to promise the blue from heaven, German for making outrageous promises) or *ins Schwarze treffen* (to hit the black, German for hit the mark), a certain part of the environment or object which has a specific color is metaphorically mapped onto a more abstract target domain. Such symbolic expressions are often only shared by speakers of the same language, as they have language as source of origin. Moreover, in these examples, color itself is only one attribute of the object central to the metaphorical expression, although an important one for determining its identity. This means that the metaphorical transfer can not take place with the color attribute alone without its frame of reference. For example, *blue* alone will not be interpreted as more surprising than other hues, and *yellow* is not interpreted as as more popular than other colors. Therefore, such metaphorical expressions are not suitable to derive cross-cultural semantic color associations that facilitate the design for intuitive use.

What is needed are expressions in which color attributes themselves provide the structure for the metaphorical transfer to more abstract target domains. In phrases like *to feel blue* (feeling depressed or discontented), *greenhorn* (being immature), *to see red* (reacting with uncontrollable rage against someone or something), or *kuroi usawa* (Japanese for black/dirty rumor), color functions as the source domain the metaphorical expression draws upon. In these examples, colors are central to the metaphorical mapping and therefore form reliable associations with the abstract concepts in the target domain. For example, the concept of immaturity is frequently associated with blue and green hues, originating from bodily and environmental experiences. Metaphorical expressions as *being blue-eyed* in English *(blauäugig* in German) link the inexperience of a caucasian baby to its eye color, which is usually blue because there is not yet a sufficient amount of color pigments in the iris that determines the later eye color[2]. Logically, this metaphor is not present in cultures with non-caucasian ethnicity. For example, in Japanese, the color metaphor *shiri ga aoi* (to have a blue buttock) exists. This metaphor links immaturity with the blue spot that appears on the back of most non-caucasian babies in the first years of life (also referred to as Mongolian spot[3]). Immaturity also has experiential links to colors in the environment. In English and German, for instance, the green horns of a young ox that have not matured yet or the green of young sprouts lead to metaphorical expressions like *greenhorn* or *grün hinter den Ohren sein* (be green behind the ears, being wet behind the ears), respectively. In Japanese, on the other hand, the bluish-green color of not yet ripe vegetables forms the basis of the metaphor *ao ni sai* (blue two years old). In the cases mentioned, environmental colors of very young animals or plants are metaphorically mapped to the more abstract concept of inexperience or immaturity as expressed in the conceptual metaphor IMMATURE IS GREEN. In other words, the physical property color as indicator of physical maturity is metaphorically extended to the domain of inexperience and naivety, although these personality traits do certainly not co-vary with color (i.e., people who act immature do not have green eye-color with a higher probability, or dress themselves in green clothes).

3.4.2.2. Metaphorical Meanings of Brightness

The color dimension *brightness* also plays an important role for conceptual scaffolding (Wells, 1910; Berlin & Kay, 1969; Valdez & Mehrabian, 1994; Crozier, 1997; Hupka et al., 1997; Gao et al., 2007). White and black are the first basic color terms in many languages across the world (Berlin & Kay, 1969) and ground a variety of abstract concepts, like VALENCE (*bright future*), INTELLIGENCE (*she is a bright girl),* HAPPINESS *(what a bright smile)* or SECRECY *(keep it dark*). As light and darkness are so fundamental experiences, some scholars argue that *bright* and *dark* itself are image schemas (Hurtienne, 2011) and therefore provide sufficient structure

[2]https://www.dwds.de/, last accessed 15.03.2017
[3]Retrieved from https://en.wikipedia.org/wiki/Mongolian_spot

to function as a source domain for metaphorical mappings. The importance of brightness probably rises from the fact that humans are diurnal animals: active at daytime and inactive during night. Much of our daily activity takes place during daytime, when our visual sense works best. A developmental perspective of how the concept of light gradually turns in a rich source concept for a variety of abstract target concepts is described by Tolaas. The author claims that as a baby, the light source is always up, so are the caring adults. Light induces a feeling of safety, because we can see our caregivers. As time goes by, light as a spatial experience is associated with the attributes of the caregivers: height, intellectual ability or moral qualities (Tolaas, 1991).

Concluding, color metaphors like those discussed within this section bridge the gap between the sensory experience of color perception and more abstract domains. Acquired very early in life based on basic physical and social interactions and reinforced through language, color metaphors are frequently encoded in and retrieved from memory, thus likely operating automatic and subconscious, also across different cultures. Although color metaphors are common in conventional speech, they are very limited in number and cannot account for the variety of abstract concepts that have to be conveyed in UIs. Because of this, this type of metaphor only plays a minor role in this work.

3.4.3. Image Schema-Color Metaphors

The second type of conceptual metaphors relevant for the deduction of cross-cultural semantic color associations are color associations with image schemas, so-called image schema-color metaphors. In these metaphors, a mapping is formed between two physical domains. Such correspondences between colors and other physical properties are expressed in *synaesthetic* metaphors like *warm yello*w or *heavy red*. The term originates from the phenomenon of synesthesia, the fusion of senses, in which a stimulation of one sensory modality leads to an involuntary experience in a second one (Dizdaroglu, Rao, Halliwell, & Gajewski, 1991).

Although sharing some superficial features, synesthesia and synaesthetic metaphors are not the same. While synesthesia operates unidirectional and cannot be explained through regular exposure, synaesthetic metaphors and underlying conceptual metaphors often operate bidirectional and arise from exposure to correlated sensory input (Deroy & Spence, 2013). Synaesthetic metaphors as surface expressions of cross-sensory conceptual mappings (Jian, 2002), called image schema-color metaphors in this work, share many features with primary metaphors (Bretones, 2001; Hampe, 2005b). Hampe defines them as "mirror[ing] properties of our perceptual apparatus itself, which interprets stimuli occupying similar positions along parallel perceptual scales (brightness, pitch, loudness) as 'correlated', hence 'similar"' (Hampe 2005, p.62). Typically, mappings occur from lower modalities onto higher than vice versa. For example, 12% of synaesthetic metaphors are of type SEEING IS TOUCHING, including mappings of touch onto color (e.g., *heavy red*) (Bretones, 2001). The other way around, mapping color onto other physical properties, is much more rare in language (e.g., *bright sound*).

Authors discussing similarities between the conceptual metaphors underlying synaesthetic expressions and primary metaphors mention their origin in perceptual correlation (Bretones, 2001; Hampe, 2005b), their ability to form the basis for other metaphors that share the same scope (Bretones, 2001), their concept-crossing and constrained nature of only mapping specific characteristics of one domain onto another (Hampe, 2005b), and that they are shared cross-linguistically (Hampe, 2005b). As such, synaesthetic metaphors might be a source of the general metaphorical capacity of the cognitive system. However, only primary metaphors cross the border between the perceptive and non-perceptive and are, as a consequence, cross-domain mappings (Hampe, 2005b). In addition, as noted by Hampe, primary metaphors are much more often used in conventional expressions compared to synaesthetic metaphors.

In image schema-color metaphors, an association between color properties and image schemas is established. Color information plays a role in the perception of those image schemas that refer to adjectives describing properties of objects and the environment (Clausner & Croft, 1999), because these are the parameters that likely co-vary with color. Examples are tactile surface properties (like SMOOTH-ROUGH) and haptic properties that result from interaction with objects (like HEAVY-LIGHT) (Ghanem, Slobodenyuk, Elhajj, Kanso, & Jraissati, 2015). A smooth surface reflects much more light than a dark surface and brighter objects appear to be lighter due to perceived object density. Although not included in the initial list of image schemas established by Johnson in *The Body in the Mind* (1987), some authors argued that certain other features fulfill the pre-conditions of image schemas, because they are frequently experienced and form the basis of many conceptual metaphors which have linguistic consequences (Grady, 1997; Lakoff, 1990; Lakoff & Johnson, 1999), and should therefore be added to the image schema inventory (Hurtienne & Blessing, 2007; Hurtienne, Stößel, & Weber, 2009; Hurtienne, 2011; Macaranas et al., 2012). In addition to the *attribute* image schemas summarized by Hurtienne 2011, the following new image schema candidates are promising for the identification of further image schema-color associations: CLEAN-DIRTY, LOUD-SILENT, OLD-YOUNG, PAINFUL-NOT PAINFUL, SMELLS GOOD-SMELLS BAD and TASTES GOOD-TASTES BAD[4]. The 16 selected image schemas that are potentially fruitful to derive cross-cultural semantic color associations that are automatically processed are summarized in Tab. 3.4.

[4]Jörn Hurtienne, personal communication, March 12, 2013

Table 3.4.: *Image Schemas that Potentially Vary with Color, Grouped by Similarity, with Linguistic Examples*

Group	Image Schema	Example Expression
Space	FAR-NEAR	distant blue
Containment	FULL-EMPTY	full red
Attribute (established)	BIG-SMALL	spacious white
	BRIGHT-DARK	dark blue
	FAST-SLOW	a quickly fading tone
	HARD-SOFT	soft beige
	HEAVY-LIGHT	heavy red
	SMOOTH-ROUGH	smooth white
	STRONG-WEAK	strong violet
	WARM-COLD	cold blue
Attribute (new)	CLEAN-DIRTY	dirty brown
	LOUD-SILENT	loud pink
	OLD-YOUNG	old gray
	PAINFUL-NOT PAINFUL	that color hurts my eyes
	SMELLS GOOD-SMELLS BAD	fresh green
	TASTES GOOD-TASTES BAD	tasty red

The origin of synaesthetic metaphors and underlying image schema-color metaphors is still under debate. While an exclusively lexical acquisition can be ruled out (Ghanem et al., 2015), authors like Hampe claim that they are either innate or early acquired in life, which is a precondition for being shared by a large amount of people and being automatically processed, thus suitable for the design for intuitive use. Investigating synaesthetic metaphors to reveal underlying image schema-color metaphors sheds light on which physical properties might be associated with certain color attributes. This information is helpful to support user interface design of physical qualities. To designers, however, more interesting than physical-to-physical mappings are physical-to-abstract mappings, which are discussed in the next section.

3.4.4. Color Substituted Image-Schematic Metaphors

The third group of conceptual metaphors that are explored regarding their potential to predict cross-cultural semantic color associations are image-schematic metaphors like INTIMATE IS WARM. In these metaphors, the *source domain* can be instantiated through color, e.g., INTIMATE IS RED. Like discussed in the previous section, colors correspond to many image schemas that structure abstract concepts through the process of metaphoric transfer. Thus, color substituted image-schematic metaphors bridge the gap between physical properties and abstract domains and cover a greater variety of abstract content than color metaphors. However, color substituted image-schematic metaphors cannot be directly accessed through conventionalized linguistic patterns of association.

In general, color is much less frequently used as *source domain* in linguistic metaphors. One reason may be its complexity and resulting proneness to errors. Color is a three-dimensional signal (hue, saturation, brightness) that represents a specific frequency of light reflected or emitted by surfaces of objects. What we perceive as color is a subjective perceptual experience constructed by our brains from input to specialized retinal cells in the eye. This means that there is considerable variance in how we see the same source of light, and the process is also error prone. Color vision deficiency is relatively common (prevalence of 8% of red-green color deficiency (protanopia or deuteranopia) in men (Gegenfurtner & Sharpe, 2001)) and severely impairs the distinguishable wavelengths. In addition, color perception is highly context-dependent. For example, in a process known under the term *color constancy*, the brain automatically adjusts the perceived color of a familiar object under varying light conditions to ensure that the color impression remains relatively constant (Krantz, 2012). This variance in color perception may also be a reason for why individuals imagine slightly different colors for a given color name, despite different approaches to standardize color names (e.g. Pantone) and the many observed regularities between different cultures and languages in categorizing basic hues (Berlin & Kay, 1969). In addition, arbitrary usage of color may partly wash out the impact of natural color associations over time (Humphrey, 1976), thereby decreasing the likelihood of being externalized in language.

Further insights for why color is less frequently used as source domain for metaphorical transfer can be gained from having a look at how color as a concept itself evolved. Mandler provides a developmental perspective on the evolution of color concepts. First of all, color is conceptualized rather late in the development (Mandler, 2008). She describes that before the age of 2-3 years, color information may be at best observed and categorized, or even remains unanalyzed perceptual information. Scholars even discuss the possibility that language is a prerequisite to acquire color concepts at all (Roberson, Davidoff, Davies, & Shapiro, 2005). When the child observes objects and events of its environment, according to Mandler, it implicitly learns similarities, resulting in the formation of perceptual schemas, in which color constitutes one element. Attentional processes are then required to extract (mainly spatial) perceptual patterns from these schemas and rescribe them as primitives that form the basis for the first concepts. "These redescriptions become associated with sensory and other bodily experiences that are not themselves redescribed, but that enrich conceptual thought." (Mandler 2008, p.207). Later, the sensory enriched primitives extend to more abstract domains and form more complex concepts. According to this view, color may fulfill the role of sensorially enhancing physical object features, which lie at the core of a variety of image schemas, and structure non-physical domains.

Regarding the mapping onto more abstract domains, Landau and colleagues describe the metaphoric relationship between source and target domain as follows: "Conceptual mappings involve systems of entailments, or mental associations between corresponding elements of the concepts in metaphoric relation. These elements can be the referents of the concepts or attributes of these referents (e.g., shape, weight, duration) as well as causal relations and other relational knowledge common to the structure of both concepts (e.g., actions known to produce seemingly spontaneous effects at a later point, such as the setting of a time bomb). Through these entailments, people are able to use select pieces of knowledge about the source concept as a structured framework for reasoning about, interpreting, and evaluating information related to the target concept." (Landau et al., 2010, p.1046). According to this, color, as a sensory enriching attribute of objects and environmental states (the referents or image schemas), constrained by the psychophysics of experiential encounters, can function as a structuring element for abstract concepts via metaphorical transfer in color substituted image-schematic metaphors. Therefore, evidence in support of this claim from research in the area of multi-sensory integration will be briefly reviewed. If color and image schemas systematically correspond in experience, there is reason to believe that color information becomes an integral part of the multi-sensory information forming image schemas, which in turn form the basis for many metaphorical mappings.

Cross-modal correspondences are "acquired, malleable, relative, and transitive pairings between sensory dimensions" (Deroy & Spence, 2013, p.643) and are shared by a large number of people (Spence, 2011). In a linguistic sense, they are synaesthetic metonymies, mirroring experiential correlations (Hampe, 2005b). Focussing on vision and touch, Ernst and Banks developed a Bayesian model that describes how the brain combines visual and haptic information from one stimulus source (Ernst & Banks, 2002). According to this model, sensory input is integrated following a maximum-likelihood estimate. For example, if we see and feel an object of a specific temperature, the visual and haptic information and resulting predictions are processed separately, weighted according to their variance and are then integrated in a statistically optimal fashion to form a more reliable estimate (Ernst, 2007; Ernst & Banks, 2002). To test their model empirically, Ernst and Banks performed an experiment in which subjects were trained to associate luminance with stiffness - two sensory signals that are not correlated in the real world. After only about two hours of training, the experimenters succeeded in inducing a link between luminance and stiffness by having changed the statistical relationship of both sensory signals. They conclude that the process and content of multi-sensory integration is not hard-wired but rather adaptive. If a new relationship between sensory signals occurs or a constant conflict between former related sensory inputs is perceived (e.g. in prism experiments), the brain recalibrates its sensory integration. The overall advantage of the process of multi-sensory integration of a correlated input from an object (*object-based perception*) or event is to increase accuracy and speed of the sensory estimate based on information from redundant sources, resulting in a rich coherent percept (Lederman, Thorne, & Jones, 1986; Stein & Meredith, 1993; Ernst & Banks, 2002; Evans & Treisman, 2010; L. Walker, Walker, & Francis, 2012; Ho, Iwai, et al., 2014). Lalanne and Lorenceau summarize the purposes cross-modal integration serves for perception and action as "including increase of salience, resolution of perceptual ambiguities, and unified perception of objects and surroundings . [..] Furthermore, the elaboration of a multimodal percept appears to be based on an adaptive combination of the contribution of each modality, according to the intrinsic reliability of sensory cue, which itself depends on the task at hand and the kind of perceptual cues involved in sensory processing." (Lalanne & Lorenceau, 2004, p.265).

Three factors are presumed to facilitate cross-modal binding: extensive spatial and temporal coincidence as well as semantic congruency (Evans & Treisman, 2010; Spence, 2011). Semantic congruency describes a match of objects or events across sensory modalities, even when they are not co-experienced on a regular basis (Woods, Spence, Butcher, & Deroy, 2013). In this case, the congruency arises from a structural or semantic overlap on conceptual dimensions like activity, evaluation and potency (Osgood et al., 1975). Consider the examples of perceiving fire and holding a hot coffee in a blue cup. In both cases, spatiotemporal congruency of color and tactile temperature cues is given and temperature awareness is conveyed through multiple complementary modalities (Balcer, 2014). However, fire as a 'naturally' colored object produces a correlated sensory stream of information in which the input from different modalities correspond, e.g., the colors red/yellow and the heat of fire. Long wavelength colors are associated with WARM and red/yellow as well as WARM have a high factor loading on the dimension of potency (Osgood et al., 1975). On the contrary, an arbitrarily colored man-made object like the blue cup filled with hot coffee may produce a sensory mismatch between different modalities. While the blue of the cup is perceived prior to the heat, expectations are based on the visual information (blue is associated with COLD), which will be disconfirmed by consecutively perceiving physical heat. In this example, the appearance of the cup generates a direct expectation of how it will feel through learned associations, cultural conventions or lawful relationships (Ludden & Kudrowitz, 2012; Ludden, Schifferstein, & Hekkert, 2009). However, color and tactile temperature cue neither match on an experiential basis (short wavelength colors naturally correlate with COLD, reinforced through cultural habits), nor on any structural dimension. This conflict or sensory incongruence is costly in terms of cognitive performance (Centerbar, Schnall, Clore, & Garvin, 2008) and results in different emotional reactions like surprise, amusement, interest, confusion or disappointment (Ludden & Kudrowitz, 2012; Ludden et al., 2009). The kind and intensity of the emotional reaction to the incongruence can be described as following an inverted u-shape (Berlyne, 1974), probably moderated through decreased liking of unfamiliar characteristics and an increased liking due to surprise reactions (Ludden, Schifferstein, & Hekkert, 2012). Thus, depending on the degree of the incongruence (specific type of *blue* and apparent physical *warmth*) the conflicting color of the cup might alter temperature perception in a contrary way compared to the tactile feedback and might cause cognitive dissonance.

Multi-sensory integration of course also occurs between modalities other than vision and touch. For example, a well-studied example of object-based perception is the constant “error” we make during eating: the brain attributes the smell of food to taste to create a whole percept (Spence, 2010). Also, there is barely an odor that is not named after the object from which it originated. In general, vision is the most important of all senses, dominating the other senses in multi-sensory integration to a degree that their input can be distorted in favor of vision (Fenko, Schifferstein, & Hekkert, 2009; Schifferstein, Fenko, Desmet, Labbe, & Martin, 2013; Schifferstein, Otten, Thoolen, & Hekkert, 2010). In their classic experiment, Rock and Victor asked participants to look at and touch an object. They created a conflict between vision and touch by distorting the visual perceived shape from the actual shape perceived by touch. As a result, the participants reported that the object felt the way it looked, suggesting that the conflict between vision and touch was completely resolved in favor of vision, and participants were unaware of the conflict (Rock & Victor, 1964). Although both visual and tactual information make object identification easiest and support memory retrieval best (Schifferstein & Cleiren, 2005; Schifferstein, 2006), vision has the advantage that it is perceived very fast and over distance and thus dominates touch in assessing materials. Without looking at materials, we are often unable to identify them correctly by touch (Wastiels, Schifferstein, Wouters, & Heylighen, 2013). Therefore, vision, as an active sense that probes information, is our primary source of objective data from the environment (Fenko, Otten, & Schifferstein, 2010). And because we 'use' and rely on our visual sense most often, a large amount of brain tissue is devoted to visual information processing (Van Essen, Felleman, DeYoe, Olavarria, & Knierim, 1990). It is also not surprising that words related to vision have a higher frequency in linguistic expressions than words related to any other sense (Viberg, 1983). Suzuki and Gyoba introduced a coefficient that indicates the extent to which words are related to sensory modalities. Even tactile sensory descriptors like *strength* received a high rating in the visual modality (Suzuki & Gyoba, 2002). More symbolic descriptors like *happy* or *familiar* were also related to sensory qualities (van Rompay, 2005; Karana, Hekkert, & Kandachar, 2007, 2009; Karana & Hekkert, 2010; Wastiels et al., 2013) and relied heavily on vision (Fenko et al., 2010). Fenko et al. thus summarize that vision occupies a salient position in the “conceptualization of the intellect” (Fenko et al., p.3317).

Color is an important visual component that plays a fundamental role in the construction of coherent percepts as well as the organization of memory (Davidoff, 1991). For example, in the robust Delboeuf illusion, a circle appears to be smaller when surrounded by a much larger circle compared to only a slightly larger circle. This illusion can not be eliminated by attentional processes or education, but is significantly reduced by decreasing the color contrast between the two circles (Van Ittersum & Wansink, 2012). Color information is independently processed (Livingstone & Hubel, 1987; Cant et al., 2008; Cavina-Pratesi, Kentridge, Heywood, & Milner, 2010) relatively early in visual analysis (Gegenfurtner, 2003) and separately stored in memory. As a surface cue, color provides important information about an object's identity (Cant et al., 2008) and facilitates object recognition. In its function as an indicator of material properties, color shapes the interaction with objects and the environment.

Concluding, people rely on implicit knowledge about cross-modal correspondences in interpreting meaning (Marks, 1990). Our brain integrates corresponding sensory input into mental representations of objects, which guide our interaction with them. Components of these representations are for example color, graspability, function or material (Lebrecht, 2012), probably represented in a multidimensional and multimodal space (Marks, 1990). Since image schemas are mental multi-modal patterns extracted from repeated perceptual and behavioral patterns, it is likely that color constitutes a visual component of them. The brief overview of literature on multi-sensory integration and cross-modal correspondences generally supports the theoretical assumption that color attributes can form an integral part of image schemas. Image schemas, in turn, constitute the *source domain* of many conceptual metaphors in image-schematic metaphors, and it might therefore be possible to predict how color is associated with more abstract domains. Color substituted image-schematic metaphors have color in the *source domain* based on experiential correspondence with the original image schema. This link is supposed to operate automatically and across cultures (sec. 3.4.3). Image-schematic metaphors are also supposed to be processed subconsciously and automatically and are shared by a large variety of people (Hurtienne, 2011). Whether these two characteristics also apply to color substituted image-schematic metaphors is an open question that will be answered by the empirical studies presented in Chapters 5-7. In the next chapter, existing empirical evidence on the three types of conceptual metaphors (color metaphors, image schema-color metaphors, and color substituted image-schematic metaphors) relevant for the predictions of cross-cultural semantic color associations is reviewed and the premises of CMToC are worked out.

4. Empirical Evidence for Conceptual Metaphor Theory of Color

In this chapter, the core premises of CMToC for predicting and explaining concrete cross-cultural semantic associations of color with physical properties and abstract concepts to facilitate the design for intuitive use are formulated. Based on the idea that metaphorical language provides hints into our mental models involving color, three different conceptual metaphors fruitful for deriving cross-cultural semantic color associations were identified: color metaphors, image schema-color metaphors, and color substituted image-schematic metaphors. The existence of these three conceptual metaphors is substantiated by reviewing empirical evidence and pointing out limitations of previous studies. The chapter closes by outlining the scope of the empirical studies.

4.1. Premises of Conceptual Metaphor Theory of Color

The fundamental tenet of CMT, developed within the field of cognitive linguistics and grounded cognition, is that a metaphor is not a mere figure of speech, but operates instead at the level of thinking (Lakoff & Johnson, 1980). Metaphors link two conceptual domains in a way that a *target domain* takes structure from a *source domain* through the metaphorical link (Deignan, 2005). The *source domain* "consists of a set of literal entities, attributes, processes and relationships, linked semantically and apparently stored together in the mind" (Deignan, 2006). The *target domain* mirrors some relationships from the *source domain*, which are "lexicalized using words and expressions from the source domain. These words and expressions are sometimes called 'linguistic metaphors' or 'metaphorical expressions' to distinguish them from conceptual metaphors." (Deignan, 2006). Thus, analyzing patterns of word use in naturally occurring language helps identifying underlying meaning, as language is a surface manifestation of thought. Applied to color psychology with the aim of predicting and explaining widely shared semantic associations between colors and physical properties as well as abstract concepts, the current work proposes the following premises of a CMToC (Conceptual Metaphor Theory of Color):

- By identifying conceptual metaphors that presumably motivate linguistic metaphors or metaphorical expressions used to talk about a certain topic, color associations of the speaker can be unveiled.

- Three types of conceptual metaphors predict cross-cultural semantic color associations with physical properties as well as abstract concepts.
  - In color metaphors, the physical *source domain* color is mapped to another, more abstract, *target domain,* e.g. SAD IS BLUE, as expressed in *I feel blue.*
  - In image schema-color metaphors, a physical *source domain* is mapped to the physical *target domain* of color, e.g. RED IS WARM, as expressed in *warm red.*
  - In image-schematic metaphors, an image schema in the *source domain* is mapped to another, more abstract, *target domain,* e.g. INTIMATE IS WARM, as expressed in *warm thoughts*. The image schema in the source domain can be replaced by a color it naturally correlates with (image schema-color metaphor), e.g. RED IS WARM, resulting in a color-substituted image-schematic metaphor, e.g. INTIMATE IS RED.
- While image schema-color metaphors indicate color associations with physical properties, color metaphors and color-substituted image-schematic metaphors denote color associations with more abstract domains. As color metaphors are very limited in number, color-substituted image-schematic metaphors are proposed to extend the available color metaphor space by image schemas. The substitution is justified as colors form an integral part of the multi-sensory information on which image schemas are based.
- Color metaphors and image schema-color metaphors can be directly derived from linguistic metaphors, but color-substituted image-schematic metaphors can not be directly derived from metaphorical expressions. Instead, they constitute a combination of image schema-color metaphor (e.g. RED IS WARM) and image-schematic metaphor (e.g. INTIMATE IS WARM), which are both expressed in linguistic metaphors.
- The origins of color associations can be categorized in innate, sensorimotor, cultural and expertise knowledge. Identifying the sources of origin of a specific color association helps to predict its cross-cultural stability and facilitates the applicability in a variety of contexts. Color associations that are innate or mandatorily acquired through basic sensorimotor experiences are, hypothetically, shared by all humans. These color associations facilitate the design of intuitive use because they are supposed to be processed automatically and subconsciously.
- Linguistic metaphors are not sufficient evidence for the existence of conceptual metaphors involving color. Empirical studies are required that test the predicted relationship between color and physical properties as well as abstract domains.

4.2. Empirical Evidence for Color Metaphors

This subchapter summarizes available experimental evidence for the five most important and well studied color metaphors.

4.2.1. Angry is Red

An example of a color metaphor that received empirical investigation is ANGER IS SEEING RED, as expressed in *seeing red* or to be *red with rage* (Lakoff, 1987). The color property hue (red) is metaphorically extended and refers to the emotional state of anger (Lakoff, 1987). Probably having its origin in physiology as the emotion of anger can result in facial flushing (Changizi et al., 2006), the ANGER IS SEEING RED metaphor is also reinforced through cultural practices across multiple cultures which express anger and danger in terms of red (Kövecses, 1990). Consequently, scholars have argued that the feeling of anger is conceptualized, at least in part, through a connection to its real-world referent red (Fetterman et al., 2011, 2012). Other conceptualizations of anger include heat and pressure (Gibbs, 1994), which are also empirically well-investigated. For example, in one study out of a series of seven experiments in total with 438 participants, Wilkowski and colleagues asked their subjects to categorize words as either anger-related or not (Wilkowski, Meier, Robinson, Carter, & Feltman, 2009). Anger-related words were processed faster when they were presented against a background showing fire compared to a background of snow. They conclude that the concept of anger is systematically linked to the cognitive representation of heat - which is in turn linked to the color dimension hue (Ho, Iwai, et al., 2014).

4.2.2. Sad is Blue

In a recent study, Reece and Danforth investigated behavioral consequences of the color metaphor SAD IS BLUE as expressed in *feeling blue.* The authors could demonstrate that by applying machine learning tools to Instagram photos it was possible to identify markers of depression in those who published them (Reece & Danforth, 2016). More depressed individuals posted photos that were more likely to be bluer, grayer and darker compared to non-depressed individuals. The link was so strong that the prediction made by this computational method outperformed in-person patient assessments by unassisted general practitioners. Thorstenson, Pazda and Elliot previously found that the link between sadness and blue might be established on a hormonal level (Thorstenson, Pazda, & Elliot, 2015). Sadness goes in hand with dopamine depletion which has an impact on retinal functioning, impairing the blue-yellow axis and overall contrast sensitivity (Bubl, Kern, Ebert, Bach, & Tebartz Van Elst, 2010). The authors found experimental evidence that perceiving color does not only affect psychological functioning, but processing emotional information influences color perception reciprocally, providing evidence for a biological underpinning of the association between color and sadness. In their experiment, subjects had to watch video clips about sadness-related vs. amusement-related or neutral content and subsequently identify different desaturated color patches. In line with metaphorical expressions like *I do see the world colorless* or *lacking brightness in mood*, the results suggest that the participants in the 'sadness' condition performed worse in the color task (i.e., identified color patches with less accuracy compared to the 'amusement' or 'neutral' condition), but only regarding the identification of color patches along the blue-yellow color axis. These results add to the slowly growing body of literature that processing emotions affect color perception like suggested by metaphorical use of color terms in conventional speech.

4.2.3. Positive is Bright - Negative is Dark

Given the ubiquity of brightness and darkness in our experiences as well as in language, many scholars have empirically studied their metaphorical extensions. The most well investigated metaphorical entailment is VALENCE: WHITE/BRIGHT IS POSITIVE - BLACK/DARK IS NEGATIVE (Hemphill, 1996; Meier et al., 2004; Naz & Epps, 2004; Smith–McLallen, Johnson, Dovidio, & Pearson, 2006; Meier et al., 2007; Sherman & Clore, 2009; Lakens, Fockenberg, Lemmens, Ham, & Midden, 2013; Meier, Fetterman, & Robinson, 2015; Sutton & Altarriba, 2015). Some authors aimed at investigating an explicit link between brightness and valence. For example, Sutton and Altarriba asked 107 participants to name the color that first came to their minds for words varying in valence. As a result, positive words were most commonly associated with the color white (Sutton & Altarriba, 2015). Other authors were more interested in unveiling implicit and automatic links between brightness and valence. For example, in a study by Meier et al., participants had to categorize the valence of words, while the brightness of the words was varied. The results showed that the categorization performance was inhibited when the font color mismatched word valence (e.g., a positive word written in black font color). However, there was no effect on response times and accuracy rates when participants had to categorize font color of valenced words, suggesting that the link between brightness and valence is asymmetric (Meier et al., 2004). Recently, Meier et al. replicated the study with a large sample size (n = 980) diverse in terms of race, age, and geographical location. Again, the response times were facilitated when word valence metaphorically matched font color (e.g., a positive word written in white font color) compared to when they did not match (e.g., a positive word written in black font color) (Meier et al., 2015).

In 2007, Meier et al. experimentally revisited the issue of the asymmetric link between brightness and valence. The participants in their study had to judge the brightness of a grayscale target after evaluating word valence of various stimuli. This time, the authors found that the perceptual judgements were biased in the direction of the metaphorically related valence in an automatic fashion (e.g., target stimuli were judged as brighter following positive evaluations) and concluded that the link between brightness and valence might be bidirectional (Meier et al., 2007). Lakens and colleagues questioned the automaticity of the WHITE IS GOOD - BLACK IS BAD association. In a series of six studies they found that while black was consistently associated with negativity, white only evoked positive associations when the negativity of black was co-activated (Lakens et al., 2012). The authors interpret the finding with the shared relational structures view: because the relationship between light and dark is similar to the relationship between positive and negative, white can be associated with positivity while the negativity of black is co-activated, but remains neutral when perceived alone (Boroditsky, 2000). Because of the context dependency of the WHITE IS GOOD association, Lakens et al. advice studying the associations of BLACK IS BAD and WHITE IS GOOD separately. A well-known and often cited study by Frank and Gilovich does examine to what extent the perception of black facilitates processing of negative concepts. The authors report that wearing black clothes in sports affect aggressiveness. Players wearing black uniforms had higher penalty records compared to players wearing uniforms in other colors. Two laboratory experiments revealed that this is due to the increased aggressiveness of the players wearing black uniforms (self-perception processes) as well as due to the biased judgements of referees seeing others wearing black uniforms (social perception) (Frank & Gilovich, 1988). A more recent study that elaborated how perceived blackness/darkness exhibits its influence on subsequent behavior is reported by Zhong et al.. In one of their experiments, the authors seated subjects in either a dimly lit or well-lit room. Those in a dimmed environment cheated more and consequently earned more undeserved money than the participants sitting in the brighter room (Zhong, Bohns, & Gino, 2010). Another study revealed that the effect of darkness on selfish-behavior is mediated through a created sense of anonymity that encourages moral transgression.

Another moderating variable in the study of brightness and its impact on cognition and behavior was recently addressed by Schietecat and colleagues. While many scholars use 'white' and 'bright' as synonyms, it seems worthwhile to distinguish between achromatic (white vs. black) and luminous brightness (bright vs. dark), as well as dynamic vs. static brightness (Schietecat, Lakens, Kort, & Ijsselsteijn, 2014). In their study the authors asked participants to rate 16 short animations of stimuli that were either static and achromatic (white, gray, black) or changed in achromatic brightness (dynamic condition, e.g. from white to black) according to how positive/negative, active/passive, powerful/weak and aggressive/calm they are perceived. Static achromatic white was rated as more positive and less aggressive than black with no differences in activity and potency. Contrasting, a rapid dynamic increase in achromatic brightness was rated as more active, aggressive and potent with no difference in valence. The authors conclude that in fast color transitions, the activity dimension becomes salient, while in static achromatic stimuli, the evaluative dimension is the most salient. Therefore, to fully exploit the benefits of street lighting on the perceived safety at night (Kort et al., 2010; Farrington & Welsh, 2008), Schietecat et al. advise to keep the detrimental effect of a too rapid increase in brightness in mind.

4.2.4. Happy is Bright - Sad is Dark

Although valence is the most well-studied metaphorical extension of brightness, the abstract domains of *happiness* and *gender* have also received considerable attention. The conceptual metaphor HAPPY IS BRIGHT - SAD IS DARK is expressed in phrases like "Lighten up!" or "He is in a dark mood" (Kövecses, 1990). Again, this metaphor is not merely a linguistic phenomenon, but has been demonstrated to be of conceptual nature, affecting subsequent behavior. For example, empirical findings in the field of emotional face recognition show that participants choose darker, cooler and more desaturated colors for sad faces (Palmer, Schloss, Xu, & Prado-León, 2013) and the processing of happy facial expressions is facilitated if the faces are shown in brighter colors (Gil & Le Bigot, 2014). This effect also goes in the opposite direction: seeing happy emotional facial expressions (smile) biases participant's perceptual judgement of facial brightness (i.e. the faces are judged to be brighter) compared to negative facial expressions (frowns) (Song, Vonasch, Meier, & Bargh, 2012). Dong et al. recently showed that feelings of hopelessness also bias people's brightness judgements (hopelessness leads to perceiving ambient lightning as darker) and discusses the potential impact on electricity consumption (Dong, Huang, & Zhong, 2015).

4.2.5. Female is Bright - Male is Dark

Finally, the gender-brightness metaphor has also received some empirical support. Due to the sexual dimorphism in skin pigmentation, females tend to have a brighter skin color than males (Hulse, 1967). In accordance with this, Semin and Palma showed that participants process male names (female names) faster when they are written in a dark font (bright font) compared to a bright font (dark font) (Semin & Palma, 2014). They apply their findings to targeted consumer products and conclude that brighter products are preferred by females, whereas darker products are preferred by men.

4.2.6. Limitations of Previous Work

So far, research has almost exclusively focused on the effects of brightness-metaphors in achromatic stimuli. One exception is the work of Hurtienne et al., 2009, who investigated the conceptual metaphors GOOD IS BRIGHT - BAD IS DARK, HAPPY IS BRIGHT - SAD IS DARK, INTELLIGENT IS BRIGHT - STUPID IS DARK, RELIGIOUS IS BRIGHT - GODLESS IS DARK, and ANGRY IS DARK with both chromatic and achromatic stimuli varying in brightness. The researchers found strong empirical support for BRIGHT-DARK mappings (except INTELLIGENT IS BRIGHT - STUPID IS DARK which was not well supported empirically) when BRIGHT-DARK was operationalized as *black* and *white*, but not when operationalized in different shades of chromatic stimuli. Whether this was a result of the specific characteristics of the stimulus material (unclear if brightness or saturation was manipulated) or a general difference between brightness in chromatic and achromatic stimuli remains a question for future research.

Color metaphors have already received considerable empirical support, probably because they come the mind most easiest. Empirical studies like those reviewed above have shown that color metaphors unveil semantic color associations that operate automatically, thus offering potential for the design for intuitive use. Color metaphors are supposed to originate from experiential correlations between colors and more abstract domains and are thus categorized into the level of sensorimotor prior knowledge. As color-to-abstract mappings, color metaphors can leverage the design of complex ideas by physical means of color. However, their number is extremely limited and therefore cannot account for the vast amount of abstract concepts that have to be conveyed in UIs.

4.3. Empirical Evidence for Image Schema-Color Metaphors

4.3.1. Literature Summary

In this section, existing scholarly work on relationships between image schemas and the color characteristics of hue, saturation, and brightness is summarized. 16 image schemas that are potentially fruitful to derive cross-cultural semantic color associations were selected based on available linguistic expressions (Tab. 3.4). Sources are spanning from HCI textbooks to questionnaires and experimental studies from various disciplines, providing a one-stop approach where evidence for color associations with sixteen image schemas can be found. In addition, the possible origins of these associations are categorized into the different levels of prior knowledge introduced in sec. 3.4.1. The section concludes with a discussion of the studies' limitations.

The result of the literature review on empirical evidence for image schema-color metaphors along with the investigated countries is presented in Tab. 4.1. Please note that a "/" in this table can mean that either no literature could be spotted by the author of this work, or the reviewed literature makes only ambiguous or no predictions at all. It does not necessarily mean that no relationship exists. The sources of origin of the associations between colors and image schemas are categorized by the author of this work into "innate", "sensorimotor", "cultural" and "expertise".

4.3.2. Limitations of Previous Work

Existing empirical work on image schema-color associations faces some limitations. First, methodological issues are discussed. Most of the extant studies are questionnaires, in which participants were tasked to assign various colors to endpoints of

Table 4.1.: *Empirical Evidence for Image Schema-Color Metaphors, Sources of Origin and Investigated Cultures*

Image Schema	Hue	Saturation	Brightness
BIG-SMALL	**innate:** increase in perceived size if highly salient longer wavelength hues or hues with high natural brightness increase background contrast, because of the physiology of the retina (2/3 of cones process longer wavelengths) (Warden & Flynn, 1931; Gundlach & Macoubrey, 1931; Wallis, 1935; Bevan & Dukes, 1953; Tedford Jr, Bergquist, & Flynn, 1977; Cleveland & McGill, 1983; Claessen, Overbeeke, & Smets, 1995; Lajos & Paris, 2009; Yoo & Smith-Jackson, 2011)	**innate:** increased perceived size if high saturation increases visual salience and background contrast (Cleveland & McGill, 1983; Abramov, Gordon, & Chan, 1991; Camgöz, Yener, & Güvenç, 2002; Xiao, Li, Luo, & Taylor, 2004; Nakano, Tanabe, Mori, Ikegami, & Fujita, 2005; Xiao, Luo, Li, Cui, & Park, 2011)	**innate:** increased perceived size if high brightness increases visual salience and background contrast (Gundlach & Macoubrey, 1931; Wallis, 1935; E. J. Robinson, 1954; Payne, 1964; Claessen et al., 1995; Camgöz et al., 2002; Kutas, Gócza, Bodrogi, & Schanda, 2004; Xiao et al., 2004; Nakano et al., 2005)
BRIGHT-DARK	**innate:** hues with high lightness are perceived brighter, because the retina is more sensitive to green-yellow hues compared to red and blue hues (Häberle, 2000; Dahm, 2006)	**innate:** intense saturation leads to an increase in brightness (Helmholtz– Kohlrausch effect) and in dim light conditions colors are perceived less saturated, because rod vision is only brightness-sensitive (Cheng, 2002; Xin, Cheng, Taylor, Sato, & Hansuebsai, 2004; Gao et al., 2007)	**innate:** positive experiential correlation between luminance and brightness (Sato, Kajiwara, Hoshino, & Nakamura, 1997b; Cheng, 2002; Xin et al., 2004; Gao et al., 2007)

Table 4.1.: *Empirical Evidence for Image Schema-Color Metaphors, Sources of Origin and Investigated Cultures (Continued)*

Image Schema	Hue	Saturation	Brightness
CLEAN-DIRTY	**sensorimotor:** gray/black/brown are the most prevalent hues of unclean matter like dust, soil, excrement and soot (Heller, 1999; Häberle, 2000; Saiz-Jimenez, 2004; Dahm, 2006)	/	**sensorimotor:** increased blackness increases perceived dirtiness due to experiential correlations (Saiz-Jimenez, 2004)
FAR-NEAR	**innate/sensorimotor:** longer wavelengths appear closer due to the retinal focus point, distant objects appear more blueish due to aerial perspective (Osgood, 1969; Verhoeff, 1941; I. L. Taylor & Sumner, 1945; Johns & Sumner, 1948; Bergum & Bergum, 1981; Egusa, 1983; Onofrei, Hunt, Siemienczuk, Touchette, & Middleton, 2004; Heller, 1999; Faubert, 1994; Häberle, 2000; Bailey, Grimm, & Davoli, 2006; Dahm, 2006)	**sensorimotor:** distant objects appear less saturated due to aerial perspective and reduced background contrast (Verhoeff, 1941; Farne, 1977; Lynch & Livingston, 2001)	**sensorimotor:** distant objects appear brighter due to aerial perspective and reduced background contrast (I. L. Taylor & Sumner, 1945; Johns & Sumner, 1948; Coules, 1955; Payne, 1964; Egusa, 1983; Lynch & Livingston, 2001)
FAST-SLOW	**innate:** reduced apparent speed of low contrast targets and increase in apparent speed of highly salient longer wavelength hues (Cavanagh, Tyler, & Favreau, 1984; L. S. Stone & Thompson, 1992; Dougherty, Press, & Wandell, 1999; Palmer et al., 2013)	**innate:** reduced apparent speed of low contrast targets (L. S. Stone & Thompson, 1992; Palmer et al., 2013)	**innate:** reduced apparent speed of low contrast targets (L. S. Stone & Thompson, 1992; Palmer et al., 2013)

Table 4.1.: *Empirical Evidence for Image Schema-Color Metaphors, Sources of Origin and Investigated Cultures (Continued)*

Image Schema	Hue	Saturation	Brightness
FULL-EMPTY	/	/	/
HARD-SOFT	**sensorimotor:** materials like metal and stone (gray, silver, black) are experientially correlated with hardness (Heller, 1999; Sato, Kajiwara, Hoshino, & Nakamura, 1997a; Häberle, 2000; Lewallen & Marini, 2003)	**sensorimotor/cultural:** lower saturation and softness are experientially correlated in certain materials (Sato et al., 1997a; Sato, Kajiwara, Hoshino, & Nakamura, 2000; Xin et al., 2004; Lucassen, Gevers, & Gijsenij, 2011; Ludwig & Simner, 2013; Yoon & Wise, 2014; Ghanem et al., 2015)	**sensorimotor/cultural:** higher brightness and softness are experientially correlated in certain materials (Cheng, 2002; Ou, Luo, Woodcock, & Wright, 2004; Gao et al., 2007; Sato et al., 2000; Xin et al., 2004; Ernst, 2007; Ludwig & Simner, 2013; Moos, Simmons, Simner, & Smith, 2013; Ghanem et al., 2015)
HEAVY-LIGHT	**sensorimotor:** hues with high relative brightness are perceived lighter due to decreased perceived object density (De Camp, 1917; C. Taylor, 1930; Pinkerton & Humphrey, 1974; Heller, 1999; Häberle, 2000; Dahm, 2006; P. Walker, Francis, & Walker, 2010)	**sensorimotor:** increased saturation increases perceived weight due to increased perceived object density (B. Wright, 1962; Alexander, 1976; Ghanem et al., 2015)	**sensorimotor:** increased blackness increases perceived weight due to increased perceived object density (Caivano, 1998; P. Walker et al., 2010; Cheng, 2002; Gao et al., 2007; Ghanem et al., 2015; Ou et al., 2004; Payne, 1964; Sato et al., 1997b; Birren, 2016; Ban et al., 2013; Sato et al., 2000; Plack & Shick, 1976; Xin et al., 2004; B. Wright, 1962; Edlin, 1969)

Table 4.1.: *Empirical Evidence for Image Schema-Color Metaphors, Sources of Origin and Investigated Cultures (Continued)*

Image Schema	Hue	Saturation	Brightness
LOUD-SILENT	**sensorimotor:** long wavelengths correspond to loud sounds due to cross-modal correspondences between multimodal sensations of intensity (Heller, 1999; Aires, 2016; Patsouras, Filippou, & Fastl, 2002; Fastl, 2004; Rader et al., 2004; Menzel, Fastl, Graf, & Hellbrück, 2008)	**sensorimotor:** high saturation corresponds to loud sounds due to cross-modal correspondences between multimodal sensations of intensity (Aires, 2016)	**sensorimotor:** high brightness corresponds to loud sounds due to cross-modal correspondences between multimodal sensations of intensity (Aires, 2016)
OLD-YOUNG	**sensorimotor:** yellow/brown/gray experientially correlate with the aging process, e.g. changes in skin and hair color (Heller, 1999; Castanet, 1997; Dahm, 2006)	**sensorimotor:** low saturation experientially correlates with the aging process, e.g. discoloration as a consequence of photodegradation during the process of aging (Castanet, 1997; Nguyen-Tri, Overbury, & Faubert, 2003; Tobin, 2008; Kaplan et al., 2011)	**sensorimotor:** darkness experientially correlates with the aging process, e.g. darkening during spoilage or as a consequence of sun exposure (Warren et al., 1991)
PAINFUL-NOT PAINFUL	**cultural:** introspectively, red and black are associated with PAINFUL and green, blue, yellow are associated with NOT PAINFUL (Huskisson, 1974; Wong & Baker, 1988; Roullet & Droulers, 2005; Moseley & Arntz, 2007; Wan, Woods, Salgado-Montejo, Velasco, & Spence, 2015)	**cultural:** introspectively, less saturation is associated with LESS PAINFUL (Buckalew & Coffield, 1982; Seifert, Speechley, Booth, Stitt, & Gibson, 1996)	/

Table 4.1.: *Empirical Evidence for Image Schema-Color Metaphors, Sources of Origin and Investigated Cultures (Continued)*

Image Schema	Hue	Saturation	Brightness
SMELLS GOOD-SMELLS BAD	**sensorimotor:** the sense of smell and taste are strongly related and hue indicates flavor identity and palatability, whereas the perceived pleasantness of an odor depends on individual preferences (Gilbert, Martin, & Kemp, 1996; Kemp & Gilbert, 1997; Gottfried & Dolan, 2003; Österbauer et al., 2005; Y. Kim, 2013)	/	**sensorimotor:** brightness is experientially correlated with less intense (and thus probably more favorable) odors (Kemp & Gilbert, 1997; Schifferstein & Tanudjaja, 2004; Y. Kim, 2013)
SMOOTH-ROUGH	**sensorimotor:** high surface reflection (like on white surfaces) is experienced on smoothness surfaces (J. Kim, Marlow, & Anderson, 2012; Ludwig & Simner, 2013; Goda, Tachibana, Okazawa, & Komatsu, 2014)	/	**sensorimotor:** high surface reflection (like on bright surfaces) is experienced on smoothness surfaces (J. Kim et al., 2012; Goda et al., 2014; Ghanem et al., 2015; Chadwick & Kentridge, 2015)
STRONG-WEAK	**sensorimotor:** long wavelength hues are associated with strength and black is associated with the strength-related concept of aggressiveness - explained by cross-modal correspondences between multimodal sensations of intensity (Adams & Osgood, 1973; Sato et al., 1997a; Green et al., 1982; Frank & Gilovich, 1988; Roullet & Droulers, 2005; Krenn, 2014)	/	**sensorimotor:** darkness is associated with strength, probably due to sexual dimorphism in skin pigmentation (Cheng, 2002; Roullet & Droulers, 2005; Gao et al., 2007)

Table 4.1.: *Empirical Evidence for Image Schema-Color Metaphors, Sources of Origin and Investigated Cultures (Continued)*

Image Schema	Hue	Saturation	Brightness
TASTES GOOD-TASTES BAD	**sensorimotor:** hue indicates flavor identity and palatability: red/orange taste better than green/yellow/brown/black/gray because of a usually higher sugar content (Duncker, 1939; Wheatley, 1973; Maga, 1974; Heller, 1999; Clydesdale, 1993; Crisosto, 1994; Bayarri, Calvo, Costell, & Durán, 2001; Spence, 2002; Lewallen & Marini, 2003; Koch & Koch, 2003; Delwiche, 2004; Dahm, 2006; Dhum, Oberfeld, Hecht, & Allendorf, 2006; Zampini, Sanabria, Phillips, & Spence, 2007; Levitan, Zampini, Li, & Spence, 2008; Shankar, Levitan, Prescott, & Spence, 2009; Harrar, Piqueras-Fiszman, & Spence, 2011; Piqueras-Fiszman, Velasco, & Spence, 2012; Piqueras-Fiszman, Alcaide, Roura, & Spence, 2012; Spence et al., 2015)	/	**sensorimotor:** brightness is associated with pleasant taste due to experiential correlations and cross-modal correspondences between multimodal sensations of intensity (Chen & Ramaswamy, 2002; Koch & Koch, 2003; Spence et al., 2015)
WARM-COLD	**sensorimotor:** long wavelength hues are experientially correlated with heat, e.g. heat sources, blushed skin (Dresslar, 1894; Tinker, 1938; Berry, 1961; B. Wright, 1962; Payne, 1964; Bennett & Rey, 1972; Morgan, Goodson, & Jones, 1975; Heller, 1999; Nakamura, Hoshino, Sato, & Kajiwara, 1996; Madden, J.T., Hewett, K., Roth, 2000; Sato et al., 2000; Häberle, 2000; Chen & Ramaswamy, 2002; Xin et al., 2004; Dahm, 2006; Ou et al., 2004; Changizi et al., 2006; Gao et al., 2007; Michael, Galich, Relland, & Prud'hon, 2010; Wastiels, Schifferstein, Heylighen, & Wouters, 2012a, 2012b; Hahn, Whitehead, Albrecht, Lefevre, & Perrett, 2012; Guéguen & Jacob, 2014; Ho, Van Doorn, et al., 2014; Ho, Iwai, et al., 2014; Buechner et al., 2015; Ho, 2015; Ghanem et al., 2015; Ziat, Balcer, Shirtz, & Rolison, 2016)	/	/

Note. /: no clear relationship found in literature semantic differentials of adjectives, to choose adjectives that represent colors best, or to pick colors that match a given adjective. This approach is suitable to unveil explicit color associations, but no statement can be made whether these associations are automatically established and have psychological consequences. Similar to general work in color psychology, studies focus on the impact of *hue* more strongly than on *saturation* and *brightness,* leaving a gap for future research. In questionnaire studies, selection criteria for color patches and adjectives are rarely provided. Rather, arbitrary stimuli are included and dimension-reducing statistical analyses are run to determine underlying basic dimensions on which adjective-color associations are evaluated. Unsurprisingly, when a factor analysis is applied to image schema-color associations (note that the term 'image schema' is not used in any study; rather, authors named the linguistic material used in their studies 'color emotion', 'color meaning', 'color image' or 'expectation'), three underlying dimensions are identified, occasionally referred to as *color heat* (hue), *color activity* (saturation) and *color weight* (brightness) (Ou et al., 2004). Such a dimension reduction does not take into account that image schemas are already indivisible basic building blocks of mental models. Moreover, many questionnaire studies do not go beyond establishing image-schema color associations. Origins and boundary conditions of color associations are seldomly discussed, and the relationship is rarely quantified, which is important to allow a transfer to other colors, concepts and contexts. These issues mainly affect the reliability of the reported evidence. If the origin and mechanism of a color association is not known, inconsistent empirical results are the consequence. Especially HCI textbooks only state general knowledge about colors and associated concepts, merging color associations from all levels of prior knowledge, thereby restricting the applicability in cross-cultural UI design. This leads to the next drawback: cultural comparisons exist, but are by far not available for all image schema-color associations. For example, while WARM-COLD is included in all cross-cultural questionnaires, CLEAN-DIRTY is rarely investigated. This imbalance in cross-cultural studies may be partly traced back to the difference between the more well-investigated established image schemas compared to the "newer". Although all image schema-color associations proposed in this work are supposed to have an innate or sensorimotor basis and are therefore mandatorily acquired, some environmental influences can shape specific associations (e.g. different colors of processed food), making cross-cultural studies indispensable.

4.4. Empirical Evidence for Color Substituted Image-Schematic Metaphors

4.4.1. Identifying Color Substituted Image-Schematic Metaphors

Color substituted image-schematic metaphors (e.g. INTIMATE IS RED) can not be directly derived from metaphorical expressions and have to be constructed from image schema-color metaphors (e.g. RED IS WARM) and image-schematic metaphors (e.g. INTIMATE IS WARM). The steps of obtaining color substituted image-schematic metaphors are:

- (1) Identifying the relationship between color and image schema as shown in Tab. 4.1.
- (2) Identifying image-schematic metaphors as metaphorical extensions of the image schema, for example through linguistic analyses and/or literature study.
- (3) Replacing the source domain in the image-schematic metaphor with the color associated with the respective image schema.

In the following paragraphs, the steps are exemplified on the image schemas BIG-SMALL, HEAVY-LIGHT, STRONG-WEAK and WARM-COLD. These four image schemas were chosen as examples because they have been shown to reliably co-vary with color and are backed up by a high amount of literature. In addition, these image schemas structure a great variety of abstract domains and therefore offer a large potential to identify relationships between color attributes and abstract concepts that would not be predicted from linguistic analyses. Moreover, these image schemas are investigated by the empirical studies in the following chapters.

4.4.1.1. Example of Big-Small

Step 1: Identification of the Image Schema-Color Relationship As summarized in Tab. 4.1, the BIG-SMALL image schema depends on foreground/background contrast and is influenced by all three color attributes. If a change in hue, saturation and/or brightness increases the contrast to the background, the perceived size is increased, thus HIGH BACKGROUND CONTRAST = BIG - SMALL BACKGROUND CONTRAST = SMALL.

Step 2: Identification of Image-Schematic Metaphors Three exemplary image-schematic metaphors involving BIG-SMALL were determined by literature research:

- KNOWING IS BIG - UNKNOWING IS SMALL (TOLAAS, 1991), as expressed in *Traveling will widen your horizon.*
- POWERFUL IS BIG - POWERLESS IS SMALL (Baldauf, 1997), as expressed in *He's a big man in industry.*
- GOOD IS BIG - BAD IS SMALL (Meier, Robinson, & Caven, 2008), as expressed in *bigger is better.*

Step 3: Replacing the Image Schema with Color As the concepts of KNOWLEDGABLE, POWER and VALENCE have now been identified as metaphorical extensions of the BIG-SMALL image schema, they can be set in relation to color. By substituting the image schema in the image-schematic metaphor with color, the color substituted image-schematic metaphors KNOWING = HIGH BACKGROUND CONTRAST - UNKNOWING = SMALL BACKGROUND CONTRAST, POWERFUL = HIGH BACKGROUND CONTRAST - POWERLESS = SMALL BACKGROUND CONTRAST and GOOD = HIGH BACKGROUND CONTRAST - BAD = SMALL BACKGROUND CONTRAST would arise. These metaphors predict that abstract concepts like KNOWLEDGABLE, POWERFUL and GOOD, are associated with a high foreground/background contrast ratio, either achieved through hue, saturation, or brightness, or a combination of these.

4.4.1.2. Example of Heavy-Light

Step 1: Identification of the Image Schema-Color Relationship As depicted in Tab. 4.1, HEAVY-LIGHT depends on hue, saturation and brightness. HUES WITH LOW RELATIVE BRIGHTNESS/HIGH SATURATION/LOW BRIGHTNESS = HEAVY - HUES WITH HIGH RELATIVE BRIGHTNESS/LOW SATURATION/HIGH BRIGHTNESS = LIGHT.

Step 2: Identification of Image-Schematic Metaphors Three exemplary image-schematic metaphors involving HEAVY-LIGHT were determined by literature research:

- IMPORTANT IS HEAVY - UNIMPORTANT IS LIGHT (Baldauf, 1997), as expressed in *heavy matters of state.*
- SAD IS HEAVY - HAPPY IS LIGHT (Stefanowitsch, 2006), as expressed in *heavy sadness.*
- SECRET IS HEAVY - TRANSPARENT IS LIGHT (Slepian, Masicampo, Toosi, & Ambady, 2012), as expressed in *heavy secrets.*

Step 3: Replacing the Image Schema with Color Substituting the image schema by color in the image-schematic metaphors determined in step 2), the color substituted image-schematic metaphors IMPORTANT = HUES WITH LOW RELATIVE BRIGHTNESS/HIGH SATURATION/LOW BRIGHTNESS - UNIMPORTANT = HUES WITH HIGH RELATIVE BRIGHTNESS/LOW SATURATION/HIGH BRIGHTNESS, SAD = HUES WITH LOW RELATIVE BRIGHTNESS/HIGH SATURATION/LOW BRIGHTNESS - HAPPY = HUES WITH HIGH RELATIVE BRIGHTNESS/LOW SATURATION/HIGH BRIGHTNESS, and SECRET = HUES WITH LOW RELATIVE BRIGHTNESS/HIGH SATURATION/LOW BRIGHTNESS - TRANSPARENT = HUES WITH HIGH RELATIVE BRIGHTNESS/LOW SATURATION/HIGH BRIGHTNESS result. These metaphors predict that people associate concepts like IMPORTANT, SAD and SECRET with dark and unsaturated hues.

4.4.1.3. Example of Strong-Weak

Step 1: Identification of the Image Schema-Color Relationship The relationship between image schema and color is taken from Tab. 4.1. The strong-weak image schema is associated with long wavelength hues such as red as well as black, and darkness further increases the link: RED/BLACK/LOW BRIGHTNESS = STRONG - HIGH BRIGHTNESS = WEAK.

Step 2: Identification of Image-Schematic Metaphors Three exemplary image-schematic metaphors involving STRONG-WEAK were determined by literature research:

- POWERFUL IS STRONG - POWERLESS IS WEAK (Baldauf, 1997), as expressed in a *strong leader*.
- COMPETENT IS STRONG - INCOMPETENT IS WEAK (Baldauf, 1997), as expressed in *Their chief strength is technology*.
- MORE IS STRONG - LESS IS WEAK (Baldauf, 1997), as expressed in *An army 20,000 strong*.

Step 3: Replacing the Image Schema with Color By replacing the image schema with color in the image-schematic metaphors identified in step 2), the following color substituted image-schematic metaphors can be derived: RED/BLACK/LOW BRIGHTNESS = POWERFUL - HIGH BRIGHTNESS = POWERLESS, RED/BLACK/LOW BRIGHTNESS = COMPETENT - HIGH BRIGHTNESS = INCOMPETENT, and RED/BLACK/LOW BRIGHTNESS = MORE - HIGH BRIGHTNESS = LESS. These metaphors predict that people associate concepts like POWERFUL, COMPETENT and MORE with dark hues, especially red and black.

4.4.1.4. Example of Warm-Cold

Step 1: Identification of the Image Schema-Color Relationship As WARM-COLD is linked to wavelength, longer wavelength hues are rated as WARM, whereas short wavelength hues are rated as COLD, see Tab. 4.1.

Step 2: Identification of Image-Schematic Metaphors Three metaphorical extensions of the WARM-COLD image schema were identified by literature research:

- INTIMATE IS WARM - DISTANT IS COLD (Lakoff & Johnson, 1999), as expressed in *We got warm with each other*.
- HAPPY IS WARM - SAD IS COLD (Stefanowitsch, 2006), as expressed in *warm joy*.
- ACTIVE IS WARM - INACTIVE IS COLD (SAMBRE, 2000), as expressed in *frozen rents*.

Step 3: Replacing the Image Schema with Color Equating the image schema with color, the following image-schematic color metaphors can be derived: INTIMATE = LONG WAVELENGTH HUES - DISTANT = SHORT WAVELENGTH HUES, HAPPY = LONG WAVELENGTH HUES - SAD = SHORT WAVELENGTH HUES, and ACTIVE = LONG WAVELENGTH HUES - PASSIVE = SHORT WAVELENGTH HUES. The metaphors predict that concepts like INTIMATE, HAPPY and ACTIVE are associated with long wavelength hues like red.

4.4.2. Empirical Studies and Limitations

Empirical studies on color substituted image-schematic metaphors are more scarce than on color metaphors or image schema-color associations, see Tab. 4.2 for an overview. In addition, these studies usually focus on the color property hue and are largely exploratory rather than theory-driven (Madden, J.T., Hewett, K., Roth, 2000).

The exemplary color substituted image-schematic metaphors systematically derived in the previous section show a considerable overlap with those reported in the literature, especially for the color attribute *hue*. For example ACTIVE IS RED, and SAD IS BLACK can be deduced from SAD = LOW BRIGHTNESS. However, with few exceptions, these works report very fixed links tied to specific colors or even hues, whereas the color substituted image-schematic color metaphors predicted by CMToC comprise a dimensional relationship. This relationship is more flexible and can be applied to different colors, thus circumventing the argumentation about specific colors with high-level symbolic meaning. Surprisingly, studies on the relationship between color and metaphorical extensions of the STRONG-WEAK image schema are largely absent, whereas color substituted image-schematic metaphors of the WARM-COLD image schema have receive considerable empirical support - a state that requires further research.

The empirical studies summarized in Tab. 4.2 rely on different methods: questionnaire studies that asked for explicit color associations vs. experimental studies that tested automatic color associations. For example, Gil and Le Bigot showed that a pink background color enhanced happy face recognition and impaired sad face recognition compared to a neutral gray background (Gil & Le Bigot, 2014). However, such experimental studies are the exception rather than the rule. The question whether color substituted image-schematic metaphors are processed automatically and are thus suitable for the design for intuitive use remains open for future research. In addition, more experimental studies are required.

Table 4.2.: *Overview of Empirically Supported Color Substituted Image-Schematic Metaphors*

Image Schema	Color-Substituted Image-Schematic Metaphors
BRIGHT-	GOOD IS WHITE/ BLUE/ GREEN (ADAMS & OSGOOD, 1973)
DARK	BAD IS BLACK/ GRAY(ADAMS & OSGOOD, 1973)
	HAPPY IS YELLOW (THARANGIE, IRFAN, MARASINGHE, & YAMADA, 2008)
	SAD IS BLACK/ BROWN (MADDEN, J.T., HEWETT, K., ROTH, 2000)
HEAVY-	SAD IS BLACK/ BROWN (MADDEN, J.T., HEWETT, K., ROTH, 2000)
LIGHT	HAPPY IS YELLOW (THARANGIE ET AL., 2008)
WARM-	ACTIVE IS RED (ADAMS & OSGOOD, 1973; MADDEN, J.T., HEWETT, K., ROTH, 2000; LEVY, 1984)
COLD	PASSIVE IS BLACK/ GRAY (ADAMS & OSGOOD, 1973), BLUE/ GREEN (MADDEN, J.T., HEWETT, K., ROTH, 2000)
	EMOTIONAL IS RED (MADDEN, J.T., HEWETT, K., ROTH, 2000)
	HAPPY IS PINK (Weller & Livingston, 1988; Gil & Le Bigot, 2014), YELLOW (THARANGIE ET AL., 2008)
	HAPPY IS BRIGHT/ SATURATED/ LONG WAVELENGTH (Dael, Perseguers, Marchand, Antonietti, & Mohr, 2016)
	SAD IS BLACK/ BROWN (MADDEN, J.T., HEWETT, K., ROTH, 2000), DARK/COOL (Palmer et al., 2013)
	INTIMATE IS RED (Baron et al., 1992; Xu & Labroo, 2014)

4.5. Research Questions and Scope of the Empirical Studies

Given the large potential of image-schema color metaphors and color substituted image-schematic metaphors for unveiling semantic color associations with physical properties as well as abstract concepts and the limitations of previous studies, a number of research questions emerges.

First, it needs to be investigated whether the 16 proposed image schema-color metaphors can be empirically supported. This is required because extant empirical work on image schema-color associations is incomplete and does not fully consider all three color attributes. Moreover, although image schema-color associations are mostly derived from experiences that are inevitably made, it is not clear whether their claimed universality can be empirically proven. This is needed to inform design decisions regarding physical information in user interfaces that can accommodate for users of different cultures.

Second, it needs to be investigated whether color substituted image-schematic metaphors can be empirically supported. As these metaphors are not expressed in linguistic metaphors, empirical evidence for their existence is needed. Also for color substituted image-schematic metaphors, the universality claim needs to be empirically proven. This is needed to inform design decisions regarding abstract information in user interfaces that can accommodate for users with different cultural backgrounds.

Third, although image schema-color metaphors and color substituted image-schematic metaphors are supposed to be frequently encoded in and retrieved from memory as form of sensorimotor prior knowledge, it is not clear whether they are processed automatically. This processing characteristic is needed to meet the demand of intuitive use.

The studies in the following chapters are dedicated to provide data towards all of the mentioned points. The first four studies are cross-cultural surveys that address the first research question by investigating the relationship between the 16 image schemas discussed above and the color attributes hue, saturation and brightness. The fifth study provides a cross-cultural experimental test of four clusters of color substituted image-schematic metaphors using measures of efficiency, effectiveness and satisfaction central to HCI (research question two). Studies 6-8 address research question three by testing whether image schema-color metaphors and color substituted image-schematic metaphors are automatically processed on the example of WARM-COLD.

5. Conveying Physical Concepts with Color

This chapter presents four online surveys that quantify the importance of the color attributes hue, saturation and brightness for 16 semantic associations with image schemas among German and Japanese subjects. The hypothesis is that specific color attributes are consistently associated with image schemas if they are regularly co-experienced in the environment. Section 5.1 presents an online survey with German participants who had to match image schemas to colors. Regression models with the three color attributes as predictors and the image schema as criterion variable were derived and the results are compared to the predictions of different theoretical frameworks. Section 5.2 replicates the same survey with Japanese subjects to assess the cultural stability of image schema-color associations. In section 5.3, the image schema-color matching task is extended to dimensional assignments of image schemas to two-color-samples to verify image schema-color associations with German participants. The same study is replicated with Japanese participants and presented in section 5.4. The chapter closes with a summary of the findings and implications for color design around physical properties in HCI.

5.1. Study 1: Color-Image Schema Associations in German Subjects

Study 1 is designed as an online survey aiming at the investigation of the relationship between 16 image schemas and the color attributes hue, saturation and brightness. While many of these image schemas stem from the haptic domain (e.g. HEAVY-LIGHT), all other senses (seeing, hearing, tasting and smelling) are covered as well (e.g. BIG-SMALL, LOUD-SILENT, TASTES GOOD - TASTES BAD, SMELLS GOOD - SMELLS BAD). In this study, German subjects rated how well certain colors match adjectives representing image schemas. Such magnitude ratings in form of Likert scales are a popular means to investigate associations (Osgood, 1969). The ratings can then be subjected to regression models that quantify the impact of the color attributes hue, saturation and brightness for each image schematic-rating, allowing to estimate the position of any given color on each image schema dimension. Next, the results can be compared against the predictions made by color-in-context theory, Affective Meanings Systems, Ecological Valence Theory on Human Color Preferences, and CMToC, which are set out in the following.

5.1.1. Predictions made by Theories in Color Psychology

5.1.1.1. Color-in-context theory

Color-in-context theory (Elliot & Maier, 2012a) states that in a given context, color perception evokes evaluative processes and motivates approach and avoidance tendencies. While the context in Study 1 might be determined by the specific rating scale, the task of performing a perceptual judgement regarding an image-schematic dimension requires more than a simple classification of the stimulus into positive or negative. Thus, the experimental paradigm is out of the scope of the theory, which focusses on approach- and avoidance behavior, and will therefore not be discussed in the results section.

5.1.1.2. Affective Meanings Systems

Affective Meanings Systems (Osgood, 1969) predicts that participants judge colors and image-schematic adjectives in a three-dimensional space of meaning (evaluation E, potency P, activity A). If the dimensional overlap in these continua is high, the adjectives and colors will be matched more likely. Tab. 5.1 summarizes the reported EPA scores of different colors and image schemas by Osgood 1973, p. 211-212, 324-325, 422 ff. (values are the averages over the 23 communities investigated in the Atlas study). Note that the EPA-loadings of other colors can not simply be interpolated as the evaluations are not equidistant, but instead need to be empirically determined.

Table 5.1.: *EPA Loadings of Different Words, after Adams and Osgood, 1973*

Word	E	P	A	Word	E	P	A	Word	E	P	A
Colors				Image Schemas				Image Schemas			
Black	-	+	-	BIG	0	+	0	SMALL	0	-	0
Blue	+	0	0	BRIGHT	+	+	+	DARK	-	0	0
Green	+	0	0	CLEAN	+	0	0	DIRTY	-	0	0
Gray	-	-	-	EMPTY	-	-	-				
Red	0	+	+	FAST	+	0	+	SLOW	0	0	-
White	+	-	0	HARD	-	+	0	SOFT	+	-	0
Yellow	0	-	0	HEAVY	0	+	-	LIGHT	+	-	0
				FAR (distant)	0	0	0	NEAR	0	0	0
				OLD	-	0	-	YOUNG	+	-	+
				PAINFUL (pain)	-	+	0				
				LOUD	-	+	+	SILENT (quiet)	+	0	-
				SMOOTH	+	-	0	ROUGH	-	+	0
				STRONG	0	+	0	WEAK	0	-	0
				WARM	+	+	+	COLD	0	0	0

Note. +: high; -: low; 0: neutral.

As Adams and Osgood only reported the different EPA loadings of colors and image schemas, the matching of both dimensions was done by the author of this work, keeping in mind that an overlap in the evaluative dimension has more impact than an overlap in potency or activity. For this, the pairs of image schemas were combined with corresponding HSB color components documented by Adams and Osgood, 1973, p.293. The results are depicted in Tab. 5.2. Whether or not an image-schematic dimension depends on color is predicted by a change in the algebraic sign for a specific color attribute between the two image-schematic end points (for a summary see also Tab. 5.8). For example, BIG is best represented by *red* and SMALL by *yellow*, and the HSB components for both do vary from + (BIG) to - (SMALL) for saturation, but are constant for hue and brightness. Therefore, it is predicted that a change in saturation will affect the image schema rating (BIG = high saturation, SMALL = low saturation), whereas hue and brightness are irrelevant based on the EPA loadings.

Table 5.2.: *Image Schemas and Corresponding Hues with HSB Component Codings taken from Adams and Osgood (1973)*

Image schema	Hue	H	S	B	Image schema	Hue	H	S	B
BIG	Red	+	+	0	SMALL	Yellow	+	-	0
BRIGHT	Blue, Green	-	+/-	0	DARK	Black, Gray	0	0	-/0
CLEAN	Blue, Green	-	+/-	0	DIRTY	Black, Gray	0	0	-/0
FAR ("distant")	Yellow	+	-	0	NEAR	Yellow	+	-	0
FAST	Blue, Green	-	+/-	0	SLOW	Yellow	+	-	0
FULL	n/a	n/a	n/a	n/a	EMPTY	Gray	0	0	0
HARD	Black	0	0	-	SOFT	White	0	0	+
HEAVY	Red	+	+	0	LIGHT	White	0	0	+
LOUD	Black	0	0	-	SILENT ("quiet")	Blue, Green	-	+/-	0
OLD	Black, Gray	0	0	-/0	YOUNG	White	0	0	+
PAINFUL ("pain")	Black	0	0	-	NOT PAINFUL	n/a	n/a	n/a	n/a
SMELLS GOOD	n/a	n/a	n/a	n/a	SMELLS BAD	n/a	n/a	n/a	n/a
SMOOTH	White	0	0	+	ROUGH	Black	0	0	-
STRONG	Red	+	+	0	WEAK	Yellow	+	-	0
TASTES GOOD	n/a	n/a	n/a	n/a	TASTES BAD	n/a	n/a	n/a	n/a
WARM	Blue, Green	-	+/-	0	COLD	Yellow	+	-	0

Note. +: red (hue)/ rich (saturation)/ bright (brightness); -: blue (hue)/ pale (saturation)/ dark (brightness); 0: neutral; n/a: no data available.

5.1.1.3. Ecological Valence Theory on Human Color Preferences

Preferred colors influence decision-making in a conscious fashion (Ho, Iwai, et al., 2014) and hue is the most important color attribute regarding color preferences (Kereney, 1966). As stated by the Ecological Valence Theory on Human Color Preferences (Palmer & Schloss, 2009), color preferences might, at least partly, originate from associations with preferred colored objects, and people tend to surround themselves with these objects. As people might systematically recall associations to preferred (or least preferred) objects, this can bias their image schema-color ratings compared to those with other color preferences. For example, when people do see a green color patch and are asked to make a BIG-SMALL judgement, they will recall positive and negative associations with big and small green colored objects. If the most dominant or plentiful associations concern big green objects, *green* will be matched with *big*. However, even people with the same favorite color are likely to differ regarding their experiential encounters with colored objects. Therefore, no precise hypotheses can be deduced from this theoretical approach without determining averaged associations with colored objects first, and linking the objects' physical properties to colors afterwards to derive semantic color associations.

5.1.1.4. Conceptual Metaphor Theory of Color

Based on CMToC, relationships between colors and image schemas are formed through experiential encounters of the fundamental statistical structures of the environment in which we live, constrained by our physical bodies. Resulting image schema-color metaphors are often expressed in metaphorical language like *warm red* or *sweet pink*. In sec. 4.3, the concrete links between hue, saturation, brightness, and 16 image schemas together with the sources of origin are summarized. The relationships have to be applied to the current experimental paradigm. As some image schemas are depending on background contrast, properties of color patches with high vs. low background contrast have first to be identified in order to make concrete predictions.

The background contrast is determined as the luminance and chrominance (hue, saturation) difference between color patch and the gray background color (HSB = n.d.,0,87) used in Study 1. The resulting foreground/background contrast after Web Content Accessibility Guidelines (WCAG) is obtained by a calculation tool provided by Hülsermann [1]. The foreground/background contrasts of the color patches used in Study 1 are depicted in Tab. 5.3. The higher the value for luminance and chrominance difference, the higher the resulting WCAG foreground/background contrast value (or contrast ratio in the form "value:1") and the higher the visibility of the color patch in the foreground. For example, according to WCAG 2.0, texts have to have at least a contrast ratio of 4.5:1 in order to be easily accessible for users. A first

[1] http://joerghuelsermann.de/tool/kontrastrechner, last accessed 16/06/2016

visual inspection of the contrast ratios in Tab. 5.3 shows that all of the bright color patches either just reach this criterion or do not fulfill it at all, and therefore have low visibility and background contrast.

Table 5.3.: *Luminance Difference (LD) and Chrominance Difference (CD) Between Color Patch and Gray Background Color (HSB = n.d.,0,87) and Resulting Foreground/Background Contrast after Web Content Accessibility Guidelines for all Stimuli in Study 1*

Name	H	S	B	Color Swatch	LD	CD	Contrast
Blue	212	35	100		30	129	1,25
	212	100	100		97	301	2,36
	212	90	25		185	498	10,83
	212	35	35		151	455	6,51
Red	0	35	100		37	169	1,49
	0	100	75		159	456	4,29
	0	100	25		190	558	9,79
	0	35	35		154	456	7,34
Yellow	56	35	100		23	137	1,31
	56	100	95		18	260	1,02
	56	100	35		130	456	4,42
	56	35	35		124	405	4,36
Orange	36	35	100		6	84	1,06
	36	100	100		62	355	1,7
	36	100	35		136	433	5,18
	36	35	35		124	405	4,36
Violet	282	35	100		154	456	7,34
	282	100	75		151	303	4,75
	282	100	25		199	558	12,88
	282	35	35		148	405	6,87
Green	130	35	90		21	170	1,03
	130	100	100		70	475	1
	130	75	25		161	560	5,42
	130	35	35		151	455	6,51
Cyan	181	35	100		1	81	1,06
	181	100	100		41	290	1,09
	181	100	35		151	455	6,51
	181	35	35		151	455	6,51
White	n.d.	0	100		35	105	1,37
Gray	n.d.	0	50		92	276	2,88
Black	n.d.	0	0		220	660	15,31

Next, the HSB values were subjected to a regression analysis with hue (dummy-coded), saturation and brightness as predictors and background contrast as criterion variable. The three color properties explain a significant amount of the variance in background contrast, $F(11, 19) = 19.913$, $p < .001$, $R^2 = .920$, *adj.* $R^2 = .874$. Details of the regression analysis are depicted in Tab. 5.4. According to these results, black has the highest background contrast and yellow the lowest. The influence of saturation closely fails to reach significance, $t(11) = 1.978$, $p = .063$, with a descriptive tendency that higher saturation increases background contrast. Finally, high brightness leads to a reduced background contrast in the stimulus material of Study 1.

Table 5.4.: *Details of the Regression Analysis of Hue, Saturation and Brightness Values on Background Contrast*

Variable	*B*	SE_B	β
Intercept	9.798	.996	
Yellow	-2.531	.941	-.230*
Blue	-0.136	.941	-.012
Cyan	-1.411	.942	-.128
Green	-2.034	.942	-.185*
Orange	-2.233	.941	-.203*
Violet	0.588	.941	.053
Black	7.074	2.066	.339*
Gray	-2.726	1.585	-.131
White	-0.044	1.606	-.002
Saturation	0.016	0.008	.190
Brightness	-0.084	0.008	-.752**

Note. B = unstandardized regression coefficient; SE_B = Standard error of the coefficient; β = standardized coefficient

Keeping the stimulus materials properties related to background contrast in mind, the following predictions are made by CMToC (see Tab. 5.5). Again, please note that a “/” in this table can mean that either no literature could be found by the author of this work, or the reviewed literature makes only ambiguous or no predictions. It does not necessarily mean that no relationship exists at all.

Table 5.5.: *Image Schemas and HSB Dependencies Predicted by Conceptual Metaphor Theory of Color in Detail*

Image Schema	Hue	Saturation	Brightness
BIG-SMALL	black is rated as the biggest color because of the highest background contrast, yellow as the smallest (e.g., BLACK IS BIG - YELLOW IS SMALL)	no or small increase in perceived size through increased saturation because of an increase in visual salience and background contrast (SATURATED IS BIG - UNSATURATED IS SMALL)	increase in perceived size through decreased brightness because of an increase in visual salience and background contrast (DARK IS BIG - BRIGHT IS SMALL)
BRIGHT-DARK	hues with high lightness are perceived brighter, because the retina is more sensitive to green-yellow hues compared to red and blue hues (e.g., YELLOW, GREEN ARE BRIGHT - BLACK, BLUE ARE DARK)	intense saturation leads to a small increase in perceived brightness (Helmholtz– Kohlrausch effect) (HIGHLY SATURATED IS BRIGHT)	positive experiential correlation between luminance and brightness (BRIGHT IS BRIGHT - DARK IS DARK)
CLEAN-DIRTY	gray/black/brown are the most prevalent hues of unclean matter like dust, soil, excrement and soot (GRAY, BLACK, BROWN ARE DIRTY)	/	increased blackness increases perceived dirtiness due to experiential correlations (BRIGHT IS CLEAN - DARK IS DIRTY)
FAR-NEAR	longer wavelength hues appear closer due to the retinal focus point, distant objects appear more blueish due to aerial perspective (e.g., BLUE IS FAR - RED IS NEAR)	distant objects appear less saturated due to aerial perspective and reduced background contrast (UNSATURATED = FAR - SATURATED = NEAR), however, arial perspective will play a minor role in this study since the viewing distance to the computer screen is short	distant objects appear brighter due to aerial perspective and reduced background contrast (BRIGHT IS FAR - DARK IS NEAR), however, arial perspective will play a minor role in this study since the viewing distance to the computer screen is short

Table 5.5.: *Image Schemas and HSB Dependencies Predicted by Conceptual Metaphor Theory of Color in Detail (Continued)*

Image Schema	Hue	Saturation	Brightness
FAST-SLOW	reduced apparent speed of low contrast targets (e.g., BLACK IS FAST - YELLOW IS SLOW)	reduced apparent speed of low contrast targets (SATURATED IS FAST - UNSATURATED IS SLOW)	reduced apparent speed of low contrast targets (DARK IS FAST - BRIGHT IS SLOW)
FULL-EMPTY	/	/	/
HARD-SOFT	materials like metal and stone (gray, silver, black) are experientially correlated with hardness (GRAY, BLACK ARE HARD)	lower saturation and softness are experientially correlated in certain materials (SATURATED IS HARD - UNSATURATED IS SOFT)	lower saturation and softness are experientially correlated in certain materials (DARK IS HARD - BRIGHT IS SOFT)
HEAVY-LIGHT	hues with high relative brightness are perceived lighter due to decreased perceived object density (e.g., BLACK, BLUE, RED ARE HEAVY - WHITE, YELLOW ARE LIGHT)	increased saturation increases perceived weight due to increased perceived object density (SATURATED IS HEAVY - UNSATURATED IS LIGHT)	increased blackness increases perceived weight due to increased perceived object density (DARK IS HEAVY - BRIGHT IS LIGHT)
LOUD-SILENT	long wavelength hues and high background contrast targets correspond to loud sounds due to cross-modal correspondences between multimodal sensations of intensity (e.g., BLACK, RED ARE LOUD - YELLOW IS SLOW)	high saturation and high background contrast targets correspond to loud sounds due to crossmodal correspondences between multimodal sensations of intensity (SATURATED IS LOUD - UNSATURATED IS SILENT)	high brightness and high background contrast targets correspond to loud sounds due to crossmodal correspondences between multimodal sensations of intensity (DARK IS LOUD - BRIGHT IS SILENT)

Table 5.5.: *Image Schemas and HSB Dependencies Predicted by Conceptual Metaphor Theory of Color in Detail (Continued)*

Image Schema	Hue	Saturation	Brightness
OLD-YOUNG	yellow/brown/gray experientially correlate with the aging process, e.g. changes in skin and hair color (YELLOW, BROWN, GRAY ARE OLD)	low saturation experientially correlates with the aging process, e.g. discoloration as a consequence of photodegradation during the process of aging (UNSATURATED IS OLD - SATURATED IS YOUNG)	darkness experientially correlates with the aging process, e.g. darkening during spoilage or as a consequence of sun exposure (DARK IS OLD - BRIGHT IS YOUNG)
PAINFUL-NOT PAINFUL	introspectively, red and black are associated with PAINFUL and green, blue, yellow are associated with NOT PAINFUL (RED, BLACK ARE PAINFUL - GREEN, BLUE. YELLOW ARE NOT PAINFUL)	introspectively, less saturation is associated with LESS PAINFUL (SATURATED IS PAINFUL - UNSATURATED IS NOT PAINFUL)	/
SMELLS GOOD-SMELLS BAD	/	/	brightness is experientially correlated with less intense (and thus probably more favorable) odors (BRIGHT SMELLS GOOD - DARK SMELLS BAD)
SMOOTH-ROUGH	high surface reflection (like on white surfaces) is experienced on smoothness surfaces (e.g., WHITE IS SMOOTH)	/	high surface reflection (like on bright surfaces) is experienced on smoothness surfaces (BRIGHT IS SMOOTH - DARK IS ROUGH)

Table 5.5.: *Image Schemas and HSB Dependencies Predicted by Conceptual Metaphor Theory of Color in Detail (Continued)*

Image Schema	Hue	Saturation	Brightness
STRONG-WEAK	long wavelength hues are associated with strength and black is associated with the strength-related concept of aggressiveness - explained by cross-modal correspondences between multimodal sensations of intensity (BLACK IS STRONG)	/	darkness is associated with strength, probably due to sexual dimorphism in skin pigmentation (DARK IS STRONG - BRIGHT IS WEAK)
TASTES GOOD-TASTES BAD	hue indicates flavor identity and palatability, experiential correlations create flavor expectations: red/orange taste better than green/yellow/brown/black/gray because of a usually higher sugar content (e.g., RED, ORANGE TASTE GOOD - GREEN, YELLOW, BROWN, BLACK, GRAY TASTE BAD)	/	brightness is associated with good taste due to experiential correlations and cross-modal correspondences between multimodal sensations of intensity (BRIGHT TASTES GOOD - DARK TASTES BAD)
WARM-COLD	long wavelength hues are experientially correlated with heat, e.g. heat sources, blushed skin (e.g., RED IS WARM - BLUE IS COLD)	/	/

Note. /: no clear relationship found in literature

5.1.2. Method

5.1.2.1. Participants

116 students were recruited from the University of Würzburg, Germany, through an online recruiting system. Students with self-reported defective color vision did not participate. All subjects had normal or corrected-to-normal vision. Data of four participants had to be excluded because they made more than one error in four plates of the Ishihara Color Test (Ishihara, 1917), which was a pre-formulated exclusion criterion for participation. The remaining 112 participants were all native German speakers (83 female, 29 male). They were between 16 and 30 years old ($M = 20.79$, $SD = 1.96$). Their favorite colors were blue (25.9%), red (17.0%), green (14.3%), black (10.7%), white (8.9%), violet (8.0%), orange (4.5%), yellow (1.8%), brown (1.8%), gray (1.8%) and others (5.4%). Participants took part in the online survey in exchange for course credit.

5.1.2.2. Procedure

The whole experiment was implemented as an online survey with two sessions conducted on different dates to relieve the burden on participants (session one: image schemas BIG-SMALL, BRIGHT-DARK, FAST-SLOW, FULL-EMPTY, HARD-SOFT, HEAVY-LIGHT, FAR-NEAR, OLD-YOUNG, SMOOTH-ROUGH, STRONG-WEAK, WARM-COLD; session two with image schemas CLEAN-DIRTY, LOUD-SILENT, PAINFUL-NOT PAINFUL, SMELLS GOOD-SMELLS BAD and TASTES GOOD-TASTES BAD). The first session was based on the list of attribute, space and containment image schemas provided by (Hurtienne, 2011). Session two contained new image schemas that are also potentially color-relevant.

In each session, participants completed four plates of the Ishihara Color Test (Ishihara, 1917) and a demographic questionnaire (see auxiliary data). They then received written instruction that their task was to rate color patches regarding different attributes ranging from one image schema pole to the other (e.g., warm-cold or bright-dark) including a neutral center (cf.sec. A.1.1). As the image schemas were not further explained to the participants, they had to rely on their intuitive understanding of them. The participants were instructed to choose the middle (neutral) answer option if they were not able to provide an evaluation for a certain attribute, i.e., image schema.

To be able to discriminate all colors more easily, participants were asked to maximize the brightness of their monitors. As the participants used their own devices, color presentation was uncontrolled. This variation was accepted to increase the ecological validity of the study. Moreover, previous research has shown that semantic judgements vary little between different presentation styles (Taft, 1997). No participant used a mobile device to complete the study.

Each color patch was evaluated concerning attributes covering the 16 image schemas of interest. The color patch was shown above the rating scales, see Fig. 5.1 for a sample screen including all sixteen image schemas). The participants could choose a specific point on the rating scale by placing the cursor above it and confirming the choice with their preferred input device (e.g. mouse, trackpad). The orders of the color patches, image schema scales and the left-right orientation of the image schema scales were randomized across participants. After the participants finished the rating part they were asked to explain how they made their choices in free text format. The survey including both sessions took about 40 minutes to complete. No outliers regarding the study completion time could be found (determined as 2.5 standard deviations above or below the average).

5.1.2.3. Material

Stimuli The online survey was created using EFS Survey. Each of the 35 shown color patches measured 1075x400 pixel (width x length) and was displayed on the top center of the screen. The questionnaire design was responsive to the display resolution. Below the color patch, sixteen seven-point attribute scales that were derived from the sixteen image schemas were displayed, e.g., warm...cold. As indicated by previous research, seven-point Likert scales are to be preferred over two-point assessments, because the more fine-grained division of the image-schematic dimension allows to express a quantitative aspect (Cheng, 2002). While 1 and 7 represented the highest association with one of the image-schematic endpoints, a rating around 4 indicated a neutral judgement. The labels were shown in 11pt Arial font against gray background (HSB = n.d.,0,87).

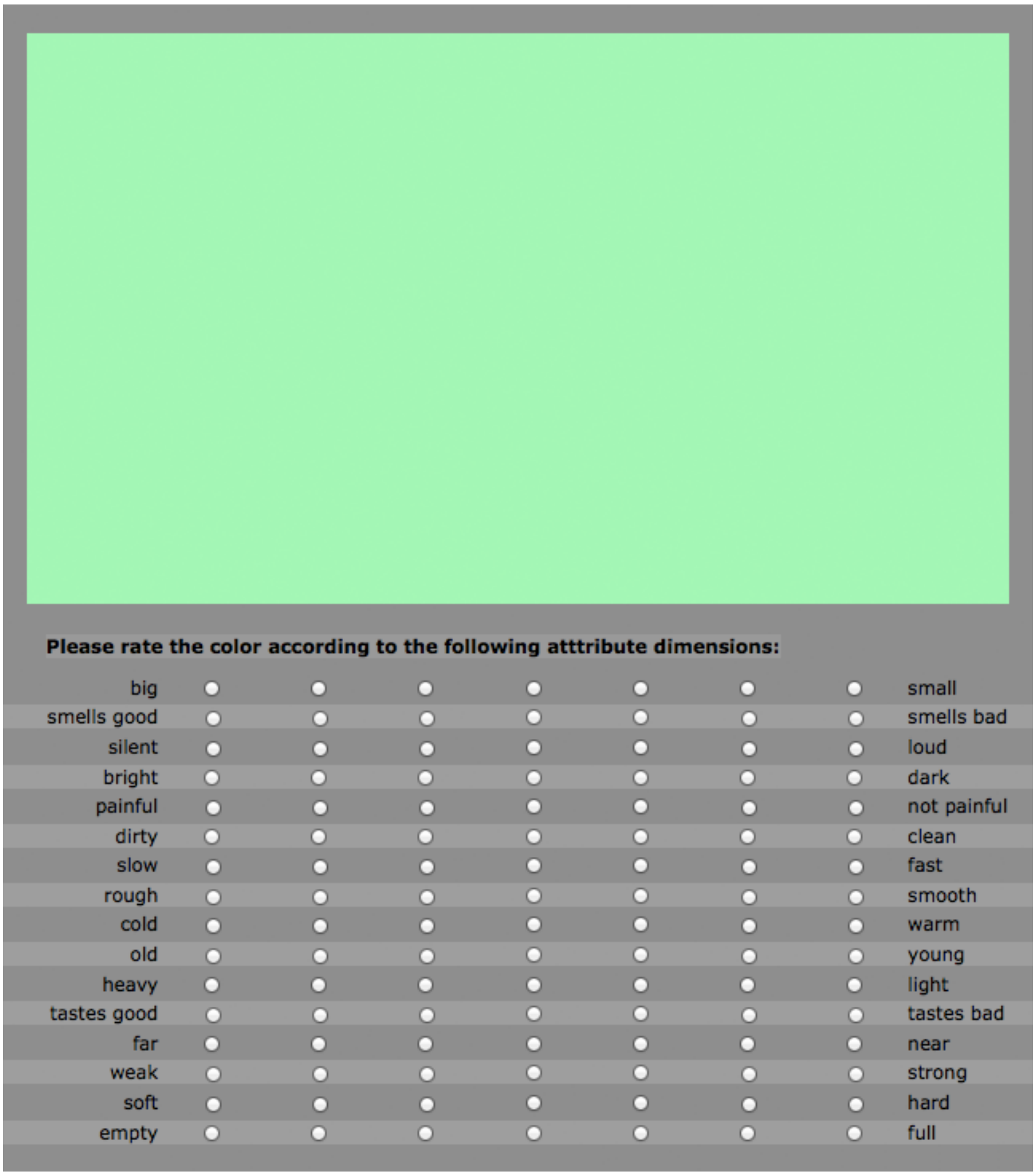

Figure 5.1.: Sample screen of Study 1 (translated from German).

The color patches were created from the primary hues blue, red and yellow, the secondary hues orange, violet and green, the tertiary hues brown and turquoise/cyan and the achromatic colors white, gray and black. This selection includes the most basic hues in most of the worlds languages red, green, yellow, blue, black and white (Varela et al., 1992) as well as additional hues that are almost never confused throughout cultures (Derefeldt, Swartling, Berggrund, & Bodrogi, 2004). For each hue, four versions were created: low saturation/high brightness, high saturation/high brightness, high saturation/low brightness and low saturation/low brightness. The high saturation value was set at 75-100%, a low saturation value at 25-35%, high brightness at 75-100% and low brightness at 25-35%. Brown was the only exception (high brightness 50%, low brightness 25%), because browns are basically darker shades of yellows and oranges. The resulting 35 color patches are depicted in Tab. A.1. The image schemas investigated in Study 1 are shown in Tab. 5.5.

Auxiliary data At the beginning of the survey, four plates of the Ishihara Color Test (Ishihara, 1917) were used to test for defective color vision. The plate numbers 1 (showing *12*), 7 (showing *74*), 8 (showing *6*) and 17 (showing *42*) were used to test for the most common forms of red-green color deficiencies. The following demographic questionnaire had to be completed with age, gender, current occupation, highest level of education, ethnicity, mother tongue and favorite color from a selection (but with the option to freely name the color if not included in the preselection).

5.1.2.4. Experimental Design and Data Analysis

A within-subjects design was used with three independent variables: hue, saturation and brightness. The dependent variable was the 7-point-rating regarding the suitability of each of the 16 bipolar image schema scales. For all studies reported in this work, data analysis was conducted using SPSS 20.0 and an effect size calculation tool by Becker (Becker, 2009). An effect size *Cohen's d* of .80 would be regarded as large, .50 as medium and .20 as small (J. Cohen, 1988). Alpha was set at .05.

While saturation and brightness are continuous scales from 0% to 100%, hue is an angular measure around the achromatic axis with red as origin (Hanbury & Serra, 2003). Because of this angular nature and the fact that only 8 specific hue values were investigated as well as the achromatic colors black, gray and white, hue was dummy-coded as categorial predictor variable before entering the analysis. The color 'brown' was excluded from this analysis, since hue, saturation and brightness could not be independently varied for this color. The hue 'red' was used as reference dummy, because it is not only defined as start and end point of the angular component hue in the the HSB color model (i.e., hues of 0 and 360 are both the same red), but is also the first named hue in most of the world's languages (Regier, Kay, & Cook, 2005).

Note that the color components hue, saturation and brightness perceived by the visual system are not entirely independent qualities (Rheingans, 1997). Perceptual brightness depends on hue as well as on saturation. If an HSB value is converted to grayscale, the individual lightness of each hue can be examined, with yellow having the highest and blue the lowest individual lightness (see Fig. 5.2). In principal, lightness is brightness with removed color information.

Brightness also depends on saturation. Compared to medium gray (0% saturation, 50% brightness), for all hues with a natural luminance above 50% (e.g. all hues in Fig. 5.2 except blue), decreasing saturation decreases brightness. Luminance is an objective measure of the intensity of the light that reaches the eye, whereas brightness is perceived luminance. If the natural luminance of a hue is below 50% (blue in Fig. 5.2), decreasing saturation increases brightness. In addition, brightness is also

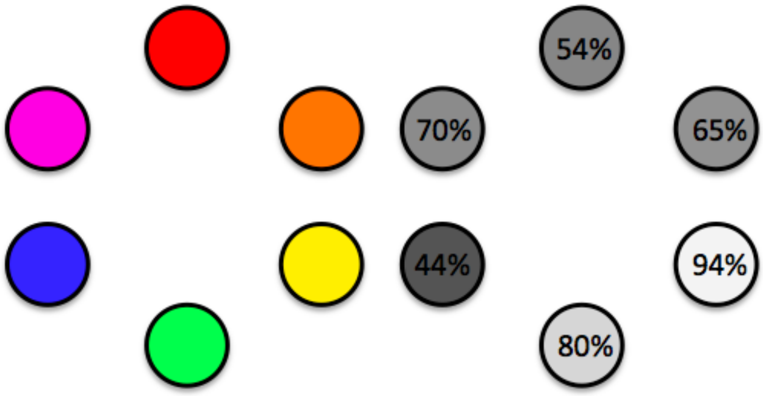

Figure 5.2.: Individual lightness of basic hues. Left: Primary and secondary hues with 100% brightness and 100% saturation. Right: Primary and secondary hues converted to grayscale with constant saturation, % of lightness compared to white (100%).

affected by surrounding colors and light conditions in the environment. However, these interdependencies between hue, saturation and brightness that distort perception in one way or the other are not strong enough to speak against including all three as predictors in the regression models (multicollinearity) (Rheingans, 1997).

5.1.3. Results

5.1.3.1. Descriptive Statistics and Regression Analysis

To examine the internal consistency of the obtained ratings with regard to the different device-dependent presentation styles, the split-half reliability between the first half and the second half of participants was calculated for each combination of image schema and color patch. The Pearson product-moment correlation coefficient indicated a strong correlation between the ratings of both halves, $r(560) = .953$, $p < .001$ (Hinkle D, Jurs S, & Wiersma W, 2003). This value is well above the recommended value of .80 suggested by Nunnally, indicating that the ratings had adequate internal reliability (Nunnally & Bernstein, 1978). The correlation coefficients for each image schema dimension as well as the means and standard deviations are documented in Tab. A.2, and Tab. A.3, respectively.

Multiple linear regression analyses were run to develop models for predicting ratings on the image schema scales from the predictors hue, saturation, and brightness, as well as to determine how much of the variation in the image schema ratings are explained by these predictor variables. The average ratings were z-standardized ($M = 0$, $SD = 1$). The assumptions of linearity, independence of errors, homoscedasticity, unusual points and normality of residuals were met. Regression coefficients, standard errors, coefficients of determination (R^2, *adjusted* R^2) and F-values for each image schema can be found in Tab. 5.6. The resulting regression models of the 16 bipolar image schema ratings are shown in Tab. A.4.

Table 5.6.: *Summary of Multiple Linear Regression Analysis for Color Variables Predicting Image Schema Ratings (N=112)*

Image Schema	Variable	B	SE_B	β	R^2	adj. R_2	$F(9,21)$
BIG-	Intercept	0.067	.149		.836***	.742	8.827
SMALL	Yellow	0.047	.140	.041			
	Blue	-0.175	.140	-.153			
	Cyan	0.045	.140	.039			
	Green	0.035	.140	.031			
	Orange	0.071	.140	.062			
	Violet	0.130	.140	.114			
	Black	-0.809	.248	-.372***			
	Gray	-0.302	.236	-.139			
	White	-0.933	.239	-.430***			
	Saturation	-0.008	.001	-.637***			
	Brightness	0.008	.001	.625***			
BRIGHT-	Intercept	1.867	.334		.947***	.897	15.026
DARK	Yellow	-0.373	.315	-.115			
	Blue	0.018	.315	.006			
	Cyan	-0.267	.316	-.083			
	Green	-0.733	.316	-.226*			
	Orange	-0.126	.315	-.039			
	Violet	-0.024	.315	-.007			
	Black	0.528	.557	.086			
	Gray	0.011	.531	.002			
	White	-0.658	.538	-.107			
	Saturation	0.006	.003	.183			
	Brightness	-0.028	.003	-.840***			
CLEAN-	Intercept	-1.336	.373		.865***	.787	11.047
DIRTY	Yellow	0.499	.262	.212			
	Blue	-0.191	.262	-.081			
	Cyan	0.113	.262	-.048			
	Green	0.049	.263	.021			
	Orange	0.557	.262	.236*			
	Violet	-0.039	.262	-.017			
	Black	-0.800	.463	-.179			
	Gray	0.473	.442	.106			
	White	-0.860	.448	-.193			
	Saturation	-0.002	.002	-.099			
	Brightness	-0.020	.002	-.847***			

Table 5.6.: *Summary of Multiple Linear Regression Analysis for Color Variables Predicting Image Schema Ratings (N=112) (Continued)*

Image Schema	Variable	B	SE_B	β	R^2	adj. R^2	$F(9,21)$
FAR-NEAR	Intercept	-0.119	.119		.845***	.755	9.399
	Yellow	-0.437	.112	-.465**			
	Blue	-0.692	.112	-.737***			
	Cyan	-0.629	.112	-.670***			
	Green	-0.337	.112	-.359**			
	Orange	-0.323	.112	-.343*			
	Violet	-0.338	.112	-.360**			
	Black	.151	.198	.085			
	Gray	-.661	.189	-.371**			
	White	-.553	.192	-.310**			
	Saturation	0.005	.001	.522***			
	Brightness	0.004	.001	.393**			
FAST-SLOW	Intercept	0.986	.193		.852***	.766	9.945
	Yellow	0.098	.182	.063			
	Blue	0.119	.182	.076			
	Cyan	-0.009	.182	-.005			
	Green	-0.131	.182	-.084			
	Orange	0.302	.182	.194			
	Violet	0.172	.182	.110			
	Black	-1.185	.321	-.400**			
	Gray	0.089	.307	.030			
	White	-0.303	.311	-.102			
	Saturation	-0.006	.002	-.437**			
	Brightness	-0.014	.002	-.867***			
FULL-EMPTY	Intercept	-0.708	.147		.894***	.833	14.607
	Yellow	0.378	.160	.235*			
	Blue	0.243	.147	.151			
	Cyan	0.359	.148	.223*			
	Green	0.121	.146	.075			
	Orange	0.271	.160	.169			
	Violet	0.073	.148	.045			
	Black	-0.346	.268	-.113			
	Gray	-0.597	.272	-.195*			
	White	-0.558	.272	-.183			
	Saturation	-0.011	.001	-.737***			
	Brightness	0.009	.001	.529***			

Table 5.6.: *Summary of Multiple Linear Regression Analysis for Color Variables Predicting Image Schema Ratings (N=112) (Continued)*

Image Schema	Variable	B	SE_{B}	β	R^2	adj. R^2	$F(9,21)$
HARD-SOFT	Intercept	-0.195	.239		.664**	.469	3.407
	Yellow	-0.142	.264	-.110			
	Blue	-0.065	.226	-.051			
	Cyan	-0.124	.226	-.096			
	Green	-0.211	.226	-.165			
	Orange	-0.010	.226	-.007			
	Violet	0.029	.226	-.023			
	Black	-1.300	.399	-.534**			
	Gray	-0.841	.380	-.345*			
	White	-0.397	.385	-.163			
	Saturation	-0.006	.002	-.474**			
	Brightness	0.006	.002	.452**			
HEAVY-LIGHT	Intercept	-1.049	.160		.963***	.941	44.480
	Yellow	0.227	.151	.088			
	Blue	0.134	.151	.052			
	Cyan	0.371	.151	.144*			
	Green	0.368	.151	.143*			
	Orange	0.008	.151	.003			
	Violet	0.105	.151	.041			
	Black	-0.106	.266	-.022			
	Gray	0.120	.254	.024			
	White	0.447	.258	.091			
	Saturation	-0.006	.001	-.231***			
	Brightness	0.024	.001	.909***			
LOUD-SILENT	Intercept	0.711	.244		.741**	.591	4.936
	Yellow	0.205	.230	.138			
	Blue	0.306	.230	.205			
	Cyan	0.283	.230	.190			
	Green	0.061	.230	.041			
	Orange	0.306	.230	.205			
	Violet	0.179	.230	.120			
	Black	-1.093	.406	-.386*			
	Gray	-0.007	.388	-.002			
	White	.0.202	.393	.071			
	Saturation	-0.012	.002	-.892***			
	Brightness	-0.002	.002	-.117			

Table 5.6.: *Summary of Multiple Linear Regression Analysis for Color Variables Predicting Image Schema Ratings (N=112) (Continued)*

Image Schema	Variable	B	SE_B	β	R^2	adj. R^2	$F(9,21)$
OLD-YOUNG	Intercept	-1.121	.160		.947***	.916	30.747
	Yellow	-0.266	.151	-.123			
	Blue	0.046	.151	.021			
	Cyan	0.202	.151	.094			
	Green	0.150	.151	.069			
	Orange	-0.437	.151	-.202**			
	Violet	0.161	.151	.075			
	Black	1.147	.266	.280***			
	Gray	-0.463	.254	-.113			
	White	-0.008	.257	-.002			
	Saturation	0.004	.001	.218**			
	Brightness	0.021	.001	.961***			
PAINFUL-NOT PAINFUL	Intercept	0.675	.243		.410	.069	1.201
	Yellow	0.036	.229	.037			
	Blue	0.324	.229	.329			
	Cyan	0.260	.229	.264			
	Green	0.077	.229	.078			
	Orange	0.136	.229	.138			
	Violet	0.256	.229	.259			
	Black	-0.547	.405	-.293			
	Gray	-0.151	.386	-.081			
	White	0.041	.391	.022			
	Saturation	-0.005	.002	-.439			
	Brightness	0.004	.002	.223			
SMELLS GOOD-SMELLS BAD	Intercept	0.711	.283		.649*	.446	3.195
	Yellow	0.495	.268	.332			
	Blue	-0.007	.267	-.005			
	Cyan	0.167	.268	.112			
	Green	0.191	.268	.128			
	Orange	0.394	.268	.264			
	Violet	-0.097	.267	-.065			
	Black	0.134	.472	.047			
	Gray	0.817	.451	.289			
	White	0.007	.457	.003			
	Saturation	-0.001	.002	-.065			
	Brightness	-0.010	.002	-.654***			

Table 5.6.: *Summary of Multiple Linear Regression Analysis for Color Variables Predicting Image Schema Ratings (N=112) (Continued)*

Image Schema	Variable	B	SE_B	β	R^2	adj. R^2	$F(9,21)$
SMOOTH-ROUGH	Intercept	0.272	.135		.926***	.884	21.712
	Yellow	0.306	.127	.198*			
	Blue	-0.005	.127	-.003			
	Cyan	0.016	.127	.010			
	Green	0.078	.127	.051			
	Orange	0.386	.127	.249**			
	Violet	-0.029	.127	-.019			
	Black	-1.296	.225	-.441***			
	Gray	0.596	.214	.203*			
	White	-0.497	.217	-.169*			
	Saturation	-0.003	.001	-.225**			
	Brightness	-0.014	.001	-.878***			
STRONG-WEAK	Intercept	-0.300	.248		.836***	.740	8.779
	Yellow	0.360	.235	.189			
	Blue	0.105	.235	.055			
	Cyan	0.324	.235	.169			
	Green	0.105	.235	.055			
	Orange	0.154	.235	.081			
	Violet	0.293	.234	.153			
	Black	-1.760	.414	-.486***			
	Gray	0.017	.395	.005			
	White	-0.922	.400	-.254			
	Saturation	-0.015	.002	-.858***			
	Brightness	0.006	.002	.319**			
TASTES GOOD-TASTES BAD	Intercept	0.740	.256		.561	.306	2.203
	Yellow	0.550	.242	.457*			
	Blue	0.212	.242	.176			
	Cyan	0.316	.242	.262			
	Green	0.286	.242	.238			
	Orange	0.413	.242	.343			
	Violet	0.210	.242	.175			
	Black	0.337	.427	.148			
	Gray	0.856	.407	.375*			
	White	0.085	.413	.037			
	Saturation	-0.002	.002	-.142			
	Brightness	-0.006	.002	-.530**			

Table 5.6.: *Summary of Multiple Linear Regression Analysis for Color Variables Predicting Image Schema Ratings (N=112) (Continued)*

Image Schema	Variable	B	SE_B	β	R^2	adj. R^2	F(9,21)
WARM-COLD	Intercept	-0.728	.254		.811***	.702	7.425
	Yellow	0.584	.240	.321*			
	Blue	1.520	.240	.835***			
	Cyan	1.415	.240	.777***			
	Green	0.961	.240	.528**			
	Orange	0.352	.240	.193			
	Violet	0.723	.239	.398**			
	Black	1.109	.423	.321*			
	Gray	1.370	.403	.397**			
	White	1.210	.409	.350**			
	Saturation	-0.004	.002	-.230			
	Brightness	-0.003	.002	-.141			

Note. B = unstandardized regression coefficient; SE_B = Standard error of the coefficient; β = standardized coefficient

For the image schema dimensions of BIG-SMALL, FAST-SLOW, FULL-EMPTY, HARD-SOFT, HEAVY-LIGHT, FAR-NEAR, OLD-YOUNG, SMOOTH-ROUGH and STRONG-WEAK, all three predictor variables (hue, saturation, brightness) explained a statistically significant proportion of the variance in the ratings. The dimension LOUD-SILENT was mainly explained by saturation and brightness. BRIGHT-DARK, CLEAN-DIRTY and TASTES GOOD-TASTES BAD were explained by hue and brightness. The remaining dimensions were explained by either hue (WARM-COLD) or brightness (SMELLS GOOD-SMELLS BAD). In case of PAINFUL-NOT PAINFUL, no statistically significant proportion of variance could be explained through the predictor variables. However, saturation almost reached statistical significance, p = .052 (see Tab. 5.7 for a summary).

Table 5.7.: *Overview of the Results of Study 1 (German Subjects)*

Image Schema	Hue- dependent	Saturation-dependent	Brightness-dependent
BIG-SMALL***	✓	✓	✓
BRIGHT-DARK***	✓	✗	✓
CLEAN-DIRTY***	✓	✗	✓
FAR-NEAR***	✓	✓	✓
FAST-SLOW***	✓	✓	✓
FULL-EMPTY***	✓	✓	✓
HARD-SOFT**	✓	✓	✓
HEAVY-LIGHT***	✓	✓	✓
LOUD-SILENT**	✓	✓	✗
OLD-YOUNG***	✓	✓	✓
PAINFUL-NOT PAINFUL	✗	✗	✗
SMELLS GOOD-SMELLS BAD*	✗	✗	✓
SMOOTH-ROUGH***	✓	✓	✓
STRONG-WEAK***	✓	✓	✓
TASTES GOOD-TASTES BAD	✓	✗	✓
WARM-COLD***	✓	✗	✗

5.1.3.2. Qualitative Data

Participants were asked in an open answer format to state how they rated the color patches regarding the image schema dimensions. Answering was optional. Answers were provided by 73 participants and were categorized into the following groups (number of mentions in brackets; multiple responses allowed):

- Gut-feeling or intuition (43), expressed in sentences like “I decided spontaneously without thinking” or “gut decision”
- Recalling associations with colored objects or experiences (33), expressed as “I had to think of objects, animals, food and so on which have these colors” or “I thought of objects and people which I associate with the colors”.
- Triggered emotions (7), as in “effects the colors have on my feelings”.

- Brightness (4), "I made my decisions based on the brightness of the colors".
- Preference (3), as in "If I liked the color or not influenced my decision".
- Other (3).

5.1.4. Discussion

The reported survey is the most comprehensive study to date with regard to the investigated amount of image schemas as well as color attributes. As a result, color-image schema regression models on the basis of hue, saturation and brightness were derived. Overall, for 14 out of 16 image schemas a significant proportion of the variance in the ratings could be explained by the predictor variables hue, saturation and brightness. Only for PAINFUL-NOT PAINFUL and TASTES GOOD-TASTES BAD no statistically significant proportion of variance could be explained through color attributes. Effect sizes (R^2) as depicted in Tab. 5.6 were consistently large (J. Cohen, 1992). The next sections discuss how well these empirical results were predicted by Affective Meanings Systems and CMToC.

5.1.4.1. Affective Meanings Systems

In total, 11 out of 48 empirical image schema-HSB relationships (22.9%) were correctly or partly correctly predicted by Affective Meanings Systems, see Tab. 5.8. These are: SATURATED IS BIG - UNSATURATED IS SMALL, GREEN IS BRIGHT, DARK IS HARD - BRIGHT IS SOFT, RED IS HEAVY, SATURATED IS HEAVY - UNSATURATED IS LIGHT, DARK IS HEAVY - BRIGHT IS LIGHT, DARK IS OLD - BRIGHT IS YOUNG, BRIGHT IS SMOOTH - DARK IS ROUGH, SATURATED IS STRONG - UNSATURATED IS WEAK, BLUE IS COLD, and WARM-COLD is independent of brightness.

Table 5.8.: *Image Schemas and HSB Dependencies Predicted by Affective Meanings Systems in Detail, Empirical Results in Parentheses for Comparison*

Image schema	Hue- dependent	Saturation- dependent	Brightness- dependent
BIG-SMALL	no ✗ (BLACK, WHITE ARE BIGGER THAN RED)	saturated - BIG, unsaturated - SMALL ✓ (SATURATED IS BIG - UNSATURATED IS SMALL)	no ✗ (DARK IS BIG - BRIGHT IS SMALL)
BRIGHT-DARK	blue, green - BRIGHT, black, gray - DARK (✓) (GREEN IS BRIGHTER THAN RED)	mixed prediction ✗ (no)	mixed prediction ✗ (BRIGHT IS BRIGHT - DARK IS DARK)
CLEAN-DIRTY	blue, green - CLEAN, black, gray - DIRTY ✗ (ORANGE IS MORE DIRTY THAN RED)	mixed prediction ✗ (no)	mixed prediction ✗ (BRIGHT IS CLEAN - DARK IS DIRTY)
FAR-NEAR	no ✗ (RED IS NEARER THAN ALL OTHER HUES EXCEPT BLACK)	no ✗ (UNSATURATED IS FAR - SATURATED IS NEAR)	no ✗ (DARK IS FAR - BRIGHT IS NEAR)
FAST-SLOW	blue, green - FAST, yellow - SLOW ✗ (BLACK IS FASTER THAN RED)	mixed prediction ✗ (SATURATED IS FAST - UNSATURATED IS SLOW)	no ✗ (BRIGHT IS FAST - DARK IS SLOW)
FULL-EMPTY	no prediction made ✗ (RED IS FULLER THAN YELLOW, BLUE, CYAN, BUT MORE EMPTY THAN GRAY)	no prediction made ✗ (SATURATED IS FULL - UNSATURATED IS EMPTY)	no prediction made ✗ (DARK IS FULL - BRIGHT IS EMPTY)
HARD-SOFT	no ✗ (BLACK, GRAY ARE HARDER THAN RED)	no ✗ (SATURATED IS HARD - UNSATURATED IS SOFT)	dark - HARD, bright - SOFT ✓ (DARK IS HARD - BRIGHT IS SOFT)

Table 5.8.: *Image Schemas and HSB Dependencies Predicted by Affective Meanings Systems in Detail, Empirical Results in Parentheses for Comparison (Continued)*

Image schema	Hue- dependent	Saturation- dependent	Brightness- dependent
HEAVY-LIGHT	red - HEAVY, white - LIGHT (✓) (RED IS HEAVIER THAN GREEN, CYAN)	saturated - HEAVY, unsaturated - LIGHT ✓ (SATURATED IS HEAVY - UNSATURATED IS LIGHT)	dark - HEAVY, bright - LIGHT ✓ (DARK IS HEAVY - BRIGHT IS LIGHT)
LOUD-SILENT (quiet)	black - LOUD, blue, green - SILENT (✓) (BLACK IS LOUDER THAN RED)	mixed prediction ✗ (SATURATED IS LOUD - UNSATURATED IS SILENT)	dark - LOUD, bright - SILENT ✗ (no)
OLD-YOUNG	no ✗ (RED IS YOUNGER THAN ORANGE AND OLDER THAN BLACK)	no ✗ (UNSATURATED IS OLD - SATURATED IS YOUNG)	dark - OLD, bright - YOUNG ✓ (DARK IS OLD - BRIGHT IS YOUNG)
PAINFUL (pain) - NOT PAINFUL	no prediction made ✗ (no)	no prediction made ✗ (no)	no prediction made ✗ (no)
SMELLS GOOD-SMELLS BAD	no prediction made ✗ (no)	no prediction made ✗ (no)	no prediction made ✗ (BRIGHT SMELLS GOOD - DARK SMELLS BAD)
SMOOTH-ROUGH	no ✗ (YELLOW, ORANGE ARE ROUGHER THAN RED - BLACK, WHITE ARE SMOOTHER THAN RED)	no ✗ (SATURATED IS SMOOTH - UNSATURATED IS ROUGH)	BRIGHT - SMOOTH, DARK - ROUGH ✓ (BRIGHT IS SMOOTH - DARK IS ROUGH)

Table 5.8.: *Image Schemas and HSB Dependencies Predicted by Affective Meanings Systems in Detail, Empirical Results in Parentheses for Comparison (Continued)*

Image schema	Hue- dependent	Saturation- dependent	Brightness- dependent
STRONG-WEAK	no ✗ (BLACK IS STRONGER THAN RED)	saturated - STRONG, unsaturated - WEAK ✓ (SATURATED IS STRONG - UNSATURATED IS WEAK)	no ✗ (DARK IS STRONG - BRIGHT IS WEAK)
TASTES GOOD-TASTES BAD	no prediction made ✗ (RED TASTES BETTER THAN YELLOW, GRAY)	no prediction made ✗ (no)	no prediction made ✗ (BRIGHT TASTES GOOD - DARK TASTES BAD)
WARM-COLD	yellow - WARM, blue, green - COLD (✓) (RED IS WARMER THAN ALL OTHER HUES EXCEPT ORANGE)	mixed predictions ✗ (no)	no ✓ (no)

Note. no: no relationship predicted; n/a: no data available. Results are marked with ✓ if the hypothesis was confirmed, and ✗if not confirmed. Results that only partly support a hypothesis are indicated by (✓).

As mentioned before, the image schemas FULL-EMPTY, PAINFUL-NOT PAINFUL, SMELLS GOOD- SMELLS BAD, and TASTES GOOD- TASTES BAD are not or not fully covered by the word material used in the original study by Osgood and colleagues (representing 11 of 48 image schema-HSB relationships). Interestingly, 8 of the 11 correctly predicted relationships are related to hue and brightness, and only three are related to the saturation component.

This points to a drawback of the current 'database' of Affective Meanings Systems. So far, there are only EPA-scores available for a small set of different hues, but not for saturation or brightness. The latter can only be indirectly accessed through comparing the change in EPA-scores of white, gray and black regarding brightness. While white-gray-black represent the dimension of achromatic brightness, chromatic brightness cannot be estimated. This situation is even more difficult for the effect of saturation, which can only be determined by comparing EPA-scores of chromatic to achromatic hues. However, as the empirical results of Study 1 showed, although white, gray and black have zero saturation compared to the other hues, they seem to represent psychologically distinct concepts, often with ratings differing from hues with low saturation. Accessing the influence of saturation only through a comparison of saturated hues vs. achromatic hues can only cover a fraction of the whole image. It is therefore not surprising that the prediction of the influence of saturated/unsaturated colors made by the Affective Meanings Systems are overall poor.

Since saturation and brightness both play an important role in many image-schematic domains, the Affective Meanings Systems database would require an extension in these dimensions. However, and this leads to another issue, the data obtained in the work by Adams and Osgood (1973) is solely based on linguistic stimuli. The color one participant imagines when reading 'red' might be different from the 'red' another participant has in mind. This would be even amplified with colors other than primary and secondary hues, such as a 'unsaturated pink'. To minimize such undesired inter-individual variation, each dimension of the EPA-score can only take one of three values: high, neutral and low. Thus, the overall measurement accuracy of EPA-scores may profit from a more fine-grained rating scale, as well as the use of less ambigue color stimuli.

5.1.4.2. Conceptual Metaphor Theory of Color

As shown in Tab. 5.8, CMToC predicted 37 out of 48 empirical image schema-HSB relationships (77.1%) correctly or partly correctly (see also Tab. 5.9 for a summary). In the following, discrepancies between the predicted and observed hue/ saturation/ brightness dependencies will be briefly discussed for each image schema.

Table 5.8.: *Image Schemas and HSB Dependencies Predicted by Conceptual Metaphor Theory of Color in Detail, Empirical Results in Parentheses for Comparison*

Image Schema	Hue-dependent	Saturation-dependent	Brightness-dependent
BIG-SMALL	BIG is high background contrast, e.g., BLACK IS BIG - YELLOW IS SMALL ✓ (BLACK, WHITE ARE BIGGER THAN RED)	BIG is high background contrast, SATURATED IS BIG - UNSATURATED IS SMALL ✓ (SATURATED IS BIG - UNSATURATED IS SMALL)	BIG is high background contrast, DARK IS BIG - BRIGHT IS SMALL ✓ (DARK IS BIG - BRIGHT IS SMALL)
BRIGHT-DARK	BRIGHT is hues with high lightness, e.g., YELLOW, GREEN ARE BRIGHT - BLACK, BLUE ARE DARK ✓ (GREEN IS BRIGHTER THAN RED)	BRIGHT is intense saturation, HIGHLY SATURATED IS BRIGHT ✗ (no)	BRIGHT is high luminance, BRIGHT IS BRIGHT - DARK IS DARK ✓ (BRIGHT IS BRIGHT - DARK IS DARK)
CLEAN-DIRTY	DIRTY is hues of unclean matter, GRAY, BLACK, BROWN ARE DIRTY ✓ (ORANGE IS MORE DIRTY THAN RED)	/ ✓ (no)	DIRTY is blackness, BRIGHT IS CLEAN - DARK IS DIRTY ✓ (BRIGHT IS CLEAN - DARK IS DIRTY)
FAR-NEAR	FAR is blueish due to aerial perspective - NEAR is longer wavelength hues due to retinal focus point, e.g., BLUE IS FAR - RED IS NEAR ✓ (RED IS NEARER THAN ALL OTHER HUES EXCEPT BLACK)	FAR is less saturated due to aerial perspective, UNSATURATED IS FAR - SATURATED IS NEAR ✓ (UNSATURATED IS FAR - SATURATED IS NEAR)	FAR is bright due to aerial perspective, BRIGHT IS FAR - DARK IS NEAR ✗ (DARK IS FAR - BRIGHT IS NEAR)
FAST-SLOW	SLOW is low contrast, e.g., BLACK, RED ARE FAST - YELLOW IS SLOW ✓ (BLACK IS FASTER THAN RED)	SLOW is low contrast, SATURATED IS FAST - UNSATURATED IS SLOW ✓ (SATURATED IS FAST - UNSATURATED IS SLOW)	slow is low contrast, DARK IS FAST - BRIGHT IS SLOW ✗ (BRIGHT IS FAST - DARK IS SLOW)

Table 5.8.: *Image Schemas and HSB Dependencies Predicted by Conceptual Metaphor Theory of Color in Detail, Empirical Results in Parentheses for Comparison (Continued)*

Image Schema	Hue-dependent	Saturation-dependent	Brightness-dependent
FULL-EMPTY	/ ✗ (RED IS FULLER THAN YELLOW, BLUE, CYAN, BUT MORE EMPTY THAN GRAY)	/ ✗ (SATURATED IS FULL - UNSATURATED IS EMPTY)	/ ✗ (DARK IS FULL - BRIGHT IS EMPTY)
HARD-SOFT	HARD is colors of hard materials, GRAY, BLACK ARE HARD ✓ (BLACK, GRAY ARE HARDER THAN RED)	SOFT is low saturation, SATURATED IS HARD - UNSATURATED IS SOFT ✓ (SATURATED IS HARD - UNSATURATED IS SOFT)	SOFT is high brightness, DARK IS HARD - BRIGHT IS SOFT ✓ (DARK IS HARD - BRIGHT IS SOFT)
HEAVY-LIGHT	LIGHT is hues with high relative brightness, e.g., BLACK, BLUE, RED ARE HEAVY - WHITE, YELLOW ARE LIGHT ✓ (RED IS HEAVIER THAN GREEN, CYAN)	HEAVY is high saturation due to perceived object density, SATURATED IS HEAVY - UNSATURATED IS LIGHT ✓ (SATURATED IS HEAVY - UNSATURATED IS LIGHT)	HEAVY is blackness due to perceived object density, DARK IS HEAVY - BRIGHT IS LIGHT ✓ (DARK IS HEAVY - BRIGHT IS LIGHT)
LOUD-SILENT	LOUD is high intensity, e.g. high background contrast or long wavelength hues due to cross-modal correspondences, e.g., BLACK, RED ARE LOUD - YELLOW IS SILENT ✓ (BLACK IS LOUDER THAN RED)	LOUD is high intensity, e.g. high background contrast or high saturation due to cross-modal correspondences, SATURATED IS LOUD - UNSATURATED IS SILENT ✓ (SATURATED IS LOUD - UNSATURATED IS SILENT)	LOUD is high intensity, e.g. high background contrast or high brightness due to cross-modal correspondences, DARK IS LOUD - BRIGHT IS SILENT ✗ (no)

Table 5.8.: *Image Schemas and HSB Dependencies Predicted by Conceptual Metaphor Theory of Color in Detail, Empirical Results in Parentheses for Comparison (Continued)*

Image Schema	Hue-dependent	Saturation-dependent	Brightness-dependent
OLD-YOUNG	age-related changes in hue, e.g. YELLOW, BROWN, GRAY ARE OLD (✓) (RED IS YOUNGER THAN ORANGE AND OLDER THAN BLACK)	OLD is low saturation due to experiential correlations, UNSATURATED IS OLD - SATURATED IS YOUNG ✓ (UNSATURATED IS OLD - SATURATED IS YOUNG)	OLD is darkness due to experiential correlations, DARK IS OLD - BRIGHT IS YOUNG) ✓ (DARK IS OLD - BRIGHT IS YOUNG)
PAINFUL-NOT PAINFUL	red and black are associated with PAINFUL and green, blue, yellow are associated with NOT PAINFUL due to experiential correlations, RED, BLACK ARE PAINFUL - GREEN, BLUE. YELLOW ARE NOT PAINFUL ✗ (no)	not painful is less saturated due to experiential correlations, SATURATED IS PAINFUL - UNSATURATED IS NOT PAINFUL ✗(no)	/ ✓ (no)
SMELLS GOOD-SMELLS BAD	/ ✓ (no)	/ ✓ (no)	SMELLS GOOD is brightness due to experiential correlations, BRIGHT SMELLS GOOD - DARK SMELLS BAD ✓ (BRIGHT SMELLS GOOD - DARK SMELLS BAD)
SMOOTH-ROUGH	SMOOTH is high surface reflection, e.g., WHITE IS SMOOTH ✓ (YELLOW, ORANGE ARE ROUGHER THAN RED - BLACK, WHITE ARE SMOOTHER THAN RED)	/ ✗ (SATURATED IS SMOOTH - UNSATURATED IS ROUGH)	SMOOTH is high surface reflection, BRIGHT IS SMOOTH - DARK IS ROUGH ✓ (BRIGHT IS SMOOTH - DARK IS ROUGH)

Table 5.8.: *Image Schemas and HSB Dependencies Predicted by Conceptual Metaphor Theory of Color in Detail, Empirical Results in Parentheses for Comparison (Continued)*

Image Schema	Hue-dependent	Saturation-dependent	Brightness-dependent
STRONG-WEAK	STRONG is long wavelength hues and black due to cross-modal correspondences, BLACK IS STRONG ✓ (BLACK IS STRONGER THAN RED)	/ ✗ (SATURATED IS STRONG - UNSATURATED IS WEAK)	STRONG is darkness due to sexual dimorphism in skin pigmentation, DARK IS STRONG - BRIGHT IS WEAK ✓ (DARK IS STRONG - BRIGHT IS WEAK)
TASTES GOOD-TASTES BAD	hue of food often differs with sugar content, e.g., RED, ORANGE TASTE GOOD - GREEN, YELLOW, BROWN, BLACK, GRAY TASTE BAD ✓ (RED TASTES BETTER THAN YELLOW, GRAY)	/ ✓ (no)	TASTES GOOD is bright due to experiential correlations and cross-modal correspondences, BRIGHT TASTES GOOD - DARK TASTES BAD ✓ (BRIGHT TASTES GOOD - DARK TASTES BAD)
WARM-COLD	WARM is longer wavelength hues due to experiential correlations, e.g., RED IS WARM - BLUE IS COLD ✓ (RED IS WARMER THAN ALL OTHER HUES EXCEPT ORANGE)	/ ✓ (no)	/ ✓ (no)

Note. /: no clear relationship found in literature. Results are marked with ✓ if the hypothesis was confirmed, and ✗if not confirmed. Results that only partly support a hypothesis are indicated by (✓).

Table 5.9.: *Overview of Empirical Image Schema-HSB Relationships Correctly Predicted by Conceptual Metaphor Theory of Color*

Image schema	HSB relationship
BIG-SMALL	BLACK IS BIG
	SATURATED IS BIG - UNSATURATED IS SMALL
	DARK IS BIG - BRIGHT IS SMALL
BRIGHT-DARK	GREEN IS BRIGHT
	BRIGHT IS BRIGHT - DARK IS DARK
CLEAN-DIRTY	ORANGE IS DIRTY
	independence of saturation
	BRIGHT IS CLEAN - DARK IS DIRTY
FAR-NEAR	RED IS NEAR
	UNSATURATED IS FAR - SATURATED IS NEAR
FAST-SLOW	BLACK IS FAST
	SATURATED IS FAST - UNSATURATED IS SLOW
HARD-SOFT	BLACK, GRAY ARE HARD
	SATURATED IS HARD - UNSATURATED IS SOFT
	DARK IS HARD - BRIGHT IS SOFT
HEAVY-LIGHT	RED IS HEAVY - GREEN, CYAN ARE LIGHT
	SATURATED IS HEAVY - UNSATURATED IS LIGHT
	DARK IS HEAVY - BRIGHT IS LIGHT
LOUD-SILENT	BLACK IS LOUD
	SATURATED IS LOUD - UNSATURATED IS SILENT
OLD-YOUNG	BROWN IS OLD
	UNSATURATED IS OLD - SATURATED IS YOUNG
	DARK IS OLD - BRIGHT IS YOUNG
PAINFUL-NOT PAINFUL	independence of hue and brightness
SMELLS GOOD-SMELLS BAD	independence of hue and saturation
	BRIGHT SMELLS GOOD - DARK SMELLS BAD
SMOOTH-ROUGH	WHITE IS SMOOTH
	BRIGHT IS SMOOTH - DARK IS ROUGH
STRONG-WEAK	BLACK IS STRONG
	DARK IS STRONG - BRIGHT IS WEAK
TASTES GOOD-TASTES BAD	RED TASTES GOOD - YELLOW, GRAY TASTE BAD
	independence of saturation
	BRIGHT TASTES GOOD - DARK TASTES BAD
WARM-COLD	RED IS WARM
	independence of saturation and brightness

bright-dark The Helmholtz–Kohlrausch effect predicts that the intense saturation of any hue is perceived as part of its luminance. However, more saturated colors were not rated as brighter than colors with lower saturation by the participants in Study

1. Although the effect hast been also shown in surface colors as those used in the study, it is more pronounced in spectral lights. In addition, the levels of saturation used in the 'high saturation' condition were not always 100%, which might have diminished the impact of saturation on BRIGHT-DARK judgements.

far-near According to the empirical data, brighter color patches were rated as nearer than darker color patches, opposite to what is predicted by the literature (distant objects appear brighter due to aerial perspective). A possible explanation is the influence of foreground/background contrast. A reduced foreground background contrast could result in increased distance ratings, as the aerial perspective reduces contrast due to the increased number of layers of air molecules with further distance (Farne, 1977). Therefore, WCAG foreground/background contrast was correlated with distance ratings, but the result did not indicate any relationship, $r(31) = .042$, $p = .822$. However, aerial perspective may only play a role outdoors and over large distances, since the amount of layers of air molecules is small over very short distances. However, in Study 1, the participants only sat about 30-60 cm away from their computer screens.

Another effect that could potentially explain the contrary results is that distance perception was influenced by perceived size. As color patches with higher background contrast are perceived bigger and the darker color patches in Study 1 had a higher foreground/background contrast ratio, it is possible that distance perception was moderated by size. By experience, closer objects are bigger than objects far away. As darker color patches appeared bigger than brighter color patches, they could be rated as closer.

fast-slow According to literature, speed perception is foreground/background contrast dependent. Apparent speed is reduced for low contrast targets. However, a Pearson correlation coefficient between participants z-standardized FAST-SLOW ratings and the WCAG foreground/background contrast values indicates a positive relationship between a high contrast ratio and slower ratings, $r(31) = .528$, $p = .002$. As the WCAG foreground/background contrast ratio is determined by both luminance and chrominance difference between foreground and background color, Pearson product-moment correlation coefficients were computed between luminance difference and WCAG contrast ratio, $r(31) = -.858$, $p < .001$, as well as between chrominance difference and WCAG contrast ratio, $p > .05$. The results show that the WCAG contrast ratio in the stimulus material used in Study 1 only correlates with brightness difference between foreground and background color, with a darker foreground color corresponding to a higher contrast ratio. Having a look at individual hues, one notices that black has the highest contrast ratio (15.31), whereas gray has a very low one (2.88). As a gray color patch on a gray background is the least visible, it is not surprising that the high contrast color patch black was rated faster than the low contrast color patch gray. For the other hues, no such pattern emerges.

A resulting research question therefore is whether chromatic and achromatic color stimuli have different effects on speed perception - also with regard to the static presentation as opposed to a dynamic presentation mode.

full-empty Regarding FULL-EMPTY, there is not yet sufficient literature to explain and predict possible associations with color. The empirical data suggests a link between hue (*red is a very full hue*), saturation (SATURATED IS FULL) as well as brightness (BRIGHT IS EMPTY). Speculating, a relationship between FULL-EMPTY and color could be moderated via object density. Hues with low relative brightness, darker and more saturated colors are perceived heavier due to an experiential association with an increase in object density. Less light is passed through opaque and dense materials, and also, for example, through containers filled with opaque liquids.

loud-silent Literature on LOUD-SILENT and color predicts that long wavelength hues, saturation and brightness increase associated LOUDNESS when they provide the most intense sensory experience. Although red received slightly higher ratings as LOUD compared to all other chromatic hues, black was rated as the loudest color. An explanation for this result could be the extreme contrast of the black color patch on the gray background, which made it stand out and easily recognizable. As black therefore has the highest intensity in terms of background contrast, it is associated with LOUD via the process of cross-modal correspondences. Also, the well-supported influence of brightness on LOUD-SILENT ratings could not be replicated in Study 1. As the bright color patches had the lowest contrast ratio with the background, it can be speculated that this obscured the effect of brightness in general. Further research is required to shed light on the influence of foreground/background contrast on judgements of LOUD-SILENT. In addition, it can be investigated if the results change by renaming the image-schematic dimensions with LOUD-SILENT synonyms that are more frequently used in language like *gaudy* (*knallig* in German) instead of LOUD when talking about colors.

painful-not painful None of the empirical relationships in Study 1 reaches significance, although the influence of higher saturation on higher ratings of PAIN almost meet the criterion, $p = .052$. This result can be interpreted as a more important effect of saturation compared to hue on pain associations and questions the sole use of hues in pain assessment tools, especially considering the high prevalence of color vision deficiency. Moreover, other means to visualize pain seem to be more important than color, e.g. size or facial expressions.

smooth-rough It was predicted that high surface reflection, like on white surfaces, is associated with SMOOTH compared to surfaces that reflect little light, like black surfaces. Although white was rated as SMOOTH, participants in Study 1 also rated black as SMOOTH. One possible explanation is that because of widespread glossy or glare-type displays which tend to reflect external light, the black color patch appeared to be glossy because of a spectral light reflection. If the contrast between specular reflection to diffuse reflection like on a black surface is high, black appears glossier than other colors with identical surface characteristics. Since the perceptual gloss of black is higher, higher ratings regarding SMOOTHNESS result (Chadwick & Kentridge, 2015).

Although not predicted by literature, a relationship between high saturation and higher ratings on SMOOTH were found. It is plausible that such a link is established because of the shared dependency between hardness, surface texture and gloss (Yoon & Wise, 2014). As many hard materials have a rather smooth surface and are more heavy (e.g., solids vs. liquids vs. gases), material density might moderate an association between surface smoothness and more saturated colors, as higher density is associated with more saturated colors. However, further research on this topic is required.

strong-weak The results of Study 1 indicate a relationship between high saturation and STRONG ratings, which is not yet supported by a considerable body of literature. Since we frequently use STRONG as a sensory descriptor for saturated colors in language (e.g. *strong red*), this relationship seems worth investigating in the future.

5.1.5. Conclusion

Color-in-context theory and Ecological Valence Theory on Human Color Preferences did not offer any concrete predictions or explanations why certain image-schematic color relationships occurred. Moreover, only 4% of the participants stated that color preferences influenced their decisions when rating the colors regarding image-schematic adjectives. Affective Meanings Systems predicted 11 out of 48 empirical image schema-HSB relationships (22.9%) correctly or partly correct. Although the predictions of Affective Meanings Systems could be extended and refined by conducting a study to determine EPA-scores for a greater variety of hues, saturation, brightness and image schemas, the majority of predictions are still mixed (e.g., saturated and unsaturated are LOUD) or not conform with the empirical data.

According to CMToC, in order to understand the cognitive relationship between image schema and color, one has to understand the physical, biological or statistical origin of the association. By following this idea, predictions can be made not only regarding the direction of the association between each color component and image schema, but also regarding the strength of the association, contextual factors and cultural stability. CMToC predicted 38 out of 48 empirical image schema-HSB relationships (77.1%) correctly or partly correctly, making it the most worthwhile theoretical approach for the prediction of semantic color associations with physical properties. While about one fifth of the predicted image schema-HSB relationships were not in line with the empirical data, CMToC still inspirits ideas on which specific situational circumstances (e.g. foreground/background contrast ratio, properties of displays, viewing distance etc.) could have led to these discrepancies, stimulating further research. The next study will have a look into cultural variations in image-schematic color associations.

5.2. Study 2: Color-Image Schema Associations in Japanese Subjects

As the first study was limited to a German participant sample, the aim of the second study is to replicate the key findings and investigate how well the derived regression models of Study 1 can predict a dataset obtained in a very different culture. Although image-schematic color associations are not likely to vary across cultures if the environmental conditions are similar, it has to be empirically proven whether these associations generalize across cultures and are consistent within them. As Europeans and Asians are said to have developed a different way of seeing and understanding the world around them for at least 2500 years (Nisbett et al., 2001; Li, Sun, & Zhang, 2007; Nisbett, 2003), Japan was chosen as a candidate culture to compare the results of the German sample with. The hypothesis is that the statistical models obtained from the survey of German participants predict a significant proportion of the observed outcomes of the same survey conducted with Japanese subjects. Subsequently, the results are discussed from the viewpoint of different theories in color psychology.

5.2.1. Method

5.2.1.1. Participants

73 students from the campus of Nihon University, Japan, volunteered to participate in the study. The participants' ages ranged from 18 to 24 years ($M = 20.6$, $SD = 1.5$). All participants (42 male, 31 female) had Japanese as their mother tongue and had no reported defective color vision. Their favorite colors were blue (17.8%), red (16.4%), orange (15.1%), white (12.3%), yellow (9.6%), violet (9.6%), green (8.2%), black (4.1%) and others (6.8%).

5.2.1.2. Procedure

The online survey was identical to Study 1 (see sec. 5.1.2 for details). After completing four plates of the Ishihara Color Test (Ishihara, 1917) and the demographic questionnaire, the subjects were instructed to judge color patches as to how well different adjectives go with a given color (see sec. A.2.1). After completing the trials, the participants were asked to explain how they made their choices. Again, the two sessions of the survey took about 40 minutes in total to complete.

5.2.1.3. Material

Stimuli The stimulus material was identical to Study 1 (sec. 5.1.2). The 16 image schemas were carefully translated into Japanese, with forward and backward check (see Tab. A.5). Each color patch was shown with either 11 (BIG-SMALL, BRIGHT-DARK, FAST-SLOW, FULL-EMPTY, HARD-SOFT, HEAVY-LIGHT, FAR-NEAR, OLD-YOUNG, SMOOTH-ROUGH, STRONG-WEAK, WARM-COLD) or five seven-item differential scales (CLEAN-DIRTY, LOUD-SILENT, PAINFUL-NOT PAINFUL, SMELLS GOOD-SMELLS BAD and TASTES GOOD-TASTES BAD) depending on the session, measuring the agreement with all 16 image schemas of interest.

Auxiliary data The same four Ishihara Color Test plates as in Study 1 were used to test for defective color vision prior to the experimental trials. The exclusion criterion was again set to more than one wrong response to the plates. The subsequent demographic questionnaire contained questions regarding age, gender, current occupation, highest level of education, ethnicity, mother tongue and favorite color.

5.2.1.4. Experimental Design and Data Analysis

The within-subjects experimental design and data analysis were identical to Study 1. First, the 7-point-rating regarding the suitability of each of the 16 image schemas was predicted by hue, saturation and brightness, coming up with regression models for the Japanese subjects. Then, using the regression models for each image schema derived from Study 1, predicted values for the Japanese subjects were computed by SPSS, rounded on three right-of-comma positions, and compared to the empirical data.

5.2.2. Results

5.2.2.1. Descriptive Statistics and Regression Analysis

The data of 73 participants that rated 35 colors regarding 16 image schema dimensions entered the analysis. The split-half reliability of the ratings was strong, $r(560) = .930$, $p < .001$, and individual responses were aggregated. The descriptive statistics (means and standard deviations) are depicted in Tab. A.6. Multiple linear regression analyses were run for each image schema rating (criterion variable) from the predictors hue, saturation, and brightness. The average ratings were z-standardized. The assumptions of linearity, independence of errors, homoscedasticity, unusual points and normality of residuals were met. Regression coefficients, standard errors, coefficients of determination (R^2, *adjusted* R^2) and F-values for each image schema can be found in Tab. A.7. The resulting regression models of the 16 bipolar image schema ratings for the Japanese subjects are shown in Tab. A.8.

For all image schemas, a statistically significant proportion of variance could be explained through the predictor variables. The image schemas BIG-SMALL, FAR-NEAR, FAST-SLOW, FULL-EMPTY, HARD-SOFT, LOUD-SILENT and STRONG-WEAK are depending on all three predictor variables hue, saturation and brightness. The dimensions CLEAN-DIRTY, HEAVY-LIGHT, OLD-YOUNG, SMELLS GOOD-SMELLS BAD, SMOOTH-ROUGH, TASTES GOOD-TASTES BAD and WARM-COLD were explained by hue and brightness. The remaining dimensions were explained by either hue (BRIGHT-DARK) or hue and saturation (PAINFUL-NOT PAINFUL), see Tab. 5.10 for a summary and comparison to the results of Study 1.

Table 5.10.: *Overview of the Results of Study 2 (Japanese Subjects) in Comparison to Study 1 (German Subjects) with an Indication of HSB-Dependency of Individual Image Schemas*

Japanese Subjects				German Subjects			
Image Schema	H	S	B	Image Schema	H	S	B
BS***	✓	✓	✓	BS***	✓	✓	✓
BD***	✓	✗	✗	BD***	✓	✗	✓
CD***	✓	✗	✓	CD***	✓	✗	✓
FN***	✓	✓	✓	FN***	✓	✓	✓
FS***	✓	✓	✓	FS***	✓	✓	✓
FE***	✓	✓	✓	FE***	✓	✓	✓
HS***	✓	✓	✓	HS**	✓	✓	✓
HL**	✓	✗	✓	HL***	✓	✓	✓
LS**	✓	✓	✓	LS**	✓	✓	✗
OY***	✓	✗	✓	OY***	✓	✓	✓
PN*	✓	✓	✗	PN	✗	✗	✗
SS***	✓	✗	✓	SS*	✗	✗	✓
SR***	✓	✗	✓	SR***	✓	✓	✓
SW***	✓	✓	✓	SW***	✓	✓	✓
TT***	✓	✗	✓	TT	✓	✗	✓
WC***	✓	✗	✓	WC***	✓	✗	✗

Note. BS (BIG-SMALL), BD (BRIGHT-DARK), CD (CLEAN-DIRTY), FN (FAR-NEAR), FS (FAST-SLOW), FE (FULL-EMPTY), HS (HARD-SOFT), HL (HEAVY-LIGHT), LS (LOUD-SILENT), OY (OLD-YOUNG), PN (PAINFUL-NOT PAINFUL), SS (SMELLS GOOD-SMELLS BAD), SR (SMOOTH-ROUGH), SW (STRONG-WEAK), TT (TASTES GOOD-TASTES BAD), WC (WARM-COLD)

Next, the observed z-standardized average ratings of the Japanese subjects were correlated across colors with the predicted values from the statistical models from Study 1. Overall, the correlation between the observed and predicted ratings was strong, $r(496) = .772$, $p < .001$. Pearson product-moment correlation coefficients for each image schema are documented in Tab. 5.11.

Table 5.11.: *Correlation Coefficients Between Observed and Predicted Ratings per Image Schema*

Image Schema	$r(31)$	p
BIG-SMALL***	.815	.000
BRIGHT-DARK***	.753	.000
CLEAN-DIRTY***	.928	.000
FAR-NEAR***	.722	.000
FAST-SLOW***	.906	.000
FULL-EMPTY***	.793	.000
HARD-SOFT**	.538	.002
HEAVY-LIGHT***	.792	.000
LOUD-SILENT**	.487	.005
OLD-YOUNG***	.894	.000
PAINFUL-NOT PAINFUL**	.567	.001
SMELLS GOOD-SMELLS BAD***	.805	.000
SMOOTH-ROUGH***	.823	.000
STRONG-WEAK***	.788	.000
TASTES GOOD-TASTES BAD***	.732	.000
WARM-COLD***	.713	.000

5.2.2.2. Qualitative Data

Participants' open answers regarding how they made their choices were clustered in the same categories as the answers of the German participants. The responses were the following (number of mentions in brackets; multiple responses allowed):

- Recalling associations with colored objects or experiences (38), as in "memories, experiences" or "experiences, like things you have seen or touched".
- Gut-feeling or intuition (23), expressed through sentences like "first impression".
- Brightness (16), as in "high and low brightness".
- Preference (5), as in "my own preferences".
- Triggered emotions (4), as in "personal feelings"
- Other (4).

5.2.3. Discussion

Study 2 was an independent test of CMToC hypotheses' with Japanese subjects. For all image schemas, a significant proportion of the variance in the ratings could be explained by the predictor variables hue, saturation and brightness. Effect sizes (R^2) as depicted in Tab. A.7 were consistently large (J. Cohen, 1992). A similar level of overall confirmation of the hypotheses could be found: 35 out of 48 empirical image schema-HSB relationships (72.9%) were correctly or partly correctly predicted (see Tab. 5.12 for a summary). As the focus of this study was on comparing the results with those obtained with German participants, discrepancies between the predicted and observed hue/ saturation/ brightness dependencies will not be discussed.

To test how well the Japanese data can be predicted from the regression models derived from the data of German participants, correlation coefficients between the predicted and actual values of the Japanese subjects were calculated. Except for LOUD-SILENT, all correlation coefficients were above .5 and represent therefore, according to Cohen, a strong correlation. While three image schemas (HARD-SOFT, LOUD-SILENT, PAINFUL-NOT PAINFUL) have correlation coefficients around .5, the other image schemas all range between .7 to .9, indicating a high overlap between the ratings of German and Japanese subjects. The subjective answers of the Japanese on how they made their decisions was also similar to those of the Germans. However, Japanese reported more often that they were influenced by recalled associations with colored objects or experience and less by gut-feeling or intuition. In addition, 'brightness' as influencing factor was reported four times as often compared to the German participants, but the results do not show more brightness-dependence in the data. However, it suggests that the Japanese subjects were more conscious about the origin of their reported image-schematic color association. One reason for this could be the different study background of the two participant samples. While the German students in Study 1 had backgrounds in human-computer systems and media communication, the Japanese students had a background in conceptual design. How do different theories in color psychology deal with issues of cultural stability and variation in semantic color associations?

Color-in-context theory argues that color meanings are grounded either in biologically based response tendencies or in learned associations that arise from frequently encountered pairings of colors with specific semantic information. Therefore, it implicitly states that as long as the learning environments between different cultures are similar, similar semantic associations will arise, while the biologically based associations are identical. Although this statement is very intuitive, the framework itself provides little help in determining the origin of different associations, therefore making not any precise predictions.

Affective Meanings Systems explicitly states that the three dimensions of affective meaning are universal and some differences in language are only superficial, grounded on a common framework of affective components on which meaning of concepts is allocated. Many empirical studies have been conducted to investigate how different

Table 5.12.: *Image Schemas and HSB Dependencies Predicted by Conceptual Metaphor Theory of Color in Detail, Empirical Results in Parentheses for Comparison*

Image Schema	Hue-dependent	Saturation-dependent	Brightness-dependent
BIG-SMALL	BIG is high background contrast, e.g., BLACK IS BIG - YELLOW IS SMALL ✓ (BLACK, WHITE ARE BIGGER THAN RED)	BIG is high background contrast, SATURATED IS BIG - UNSATURATED IS SMALL ✓ (SATURATED IS BIG - UNSATURATED IS SMALL)	BIG is high background contrast, DARK IS BIG - BRIGHT IS SMALL ✓ (DARK IS BIG - BRIGHT IS SMALL)
BRIGHT-DARK	BRIGHT is hues with high relative brightness, e.g., YELLOW, GREEN ARE BRIGHT - BLACK, BLUE ARE DARK ✗ (no)	BRIGHT is intense saturation, HIGHLY SATURATED IS BRIGHT ✗ (no)	BRIGHT is high luminance, BRIGHT IS BRIGHT - DARK IS DARK ✓ (BRIGHT IS BRIGHT - DARK IS DARK)
CLEAN-DIRTY	DIRTY is hues of unclean matter, GRAY, BLACK, BROWN ARE DIRTY ✓ (ORANGE IS MORE DIRTY THAN RED)	/ ✓ (no)	DIRTY is blackness, BRIGHT IS CLEAN - DARK IS DIRTY ✓ (BRIGHT IS CLEAN - DARK IS DIRTY)
FAR-NEAR	FAR is blueish due to aerial perspective - NEAR is longer wavelength hues due to retinal focus point, e.g., BLUE IS FAR - RED IS NEAR ✓ (RED IS NEARER THAN BLUE, CYAN, VIOLET AND GRAY)	FAR is less saturated due to aerial perspective, UNSATURATED IS FAR - SATURATED IS NEAR ✓ (UNSATURATED IS FAR - SATURATED IS NEAR)	FAR is bright due to aerial perspective, BRIGHT IS FAR - DARK IS NEAR ✗ (DARK IS FAR - BRIGHT IS NEAR)
FAST-SLOW	SLOW is low contrast, e.g., BLACK, RED ARE FAST - YELLOW IS SLOW ✓ (BLACK AND WHITE ARE FASTER THAN RED)	SLOW is low contrast, SATURATED IS FAST - UNSATURATED IS SLOW ✓ (SATURATED IS FAST - UNSATURATED IS SLOW)	slow is low contrast, DARK IS FAST - BRIGHT IS SLOW ✗ (BRIGHT IS FAST - DARK IS SLOW)

Table 5.12.: *Image Schemas and HSB Dependencies Predicted by Conceptual Metaphor Theory of Color in Detail, Empirical Results in Parentheses for Comparison (Continued)*

Image Schema	Hue-dependent	Saturation-dependent	Brightness-dependent
FULL-EMPTY	/ ✗ (RED IS FULLER THAN BLUE, CYAN)	/ ✗ (SATURATED IS FULL - UNSATURATED IS EMPTY)	/ ✗ (DARK IS FULL - BRIGHT IS EMPTY)
HARD-SOFT	HARD is colors of hard materials, GRAY, BLACK ARE HARD ✓ (BLACK, GRAY ARE HARDER THAN RED)	SOFT is low saturation, SATURATED IS HARD - UNSATURATED IS SOFT ✓ (SATURATED IS HARD - UNSATURATED IS SOFT)	SOFT is high brightness, DARK IS HARD - BRIGHT IS SOFT ✓ (DARK IS HARD - BRIGHT IS SOFT)
HEAVY-LIGHT	LIGHT is hues with high relative brightness, e.g., BLACK, BLUE, RED ARE HEAVY - WHITE, YELLOW ARE LIGHT ✓ (BLUE IS HEAVIER THAN RED)	HEAVY is high saturation due to perceived object density, SATURATED IS HEAVY - UNSATURATED IS LIGHT ✓ (SATURATED IS HEAVY - UNSATURATED IS LIGHT)	HEAVY is blackness due to perceived object density, DARK IS HEAVY - BRIGHT IS LIGHT ✓ (DARK IS HEAVY - BRIGHT IS LIGHT)
LOUD-SILENT	LOUD is high intensity, e.g. high background contrast or long wavelength hues due to cross-modal correspondences, e.g., BLACK, RED ARE LOUD - YELLOW IS SILENT ✓ (RED IS LOUDER THAN BLUE, CYAN, WHITE)	LOUD is high intensity, e.g. high background contrast or high saturation due to cross-modal correspondences, SATURATED IS LOUD - UNSATURATED IS SILENT ✓ (SATURATED IS LOUD - UNSATURATED IS SILENT)	LOUD is high intensity, e.g. high background contrast or high brightness due to cross-modal correspondences, DARK IS LOUD - BRIGHT IS SILENT ✗ (BRIGHT IS LOUD - DARK IS SILENT)
OLD-YOUNG	age-related changes in hue, e.g. YELLOW, BROWN, GRAY ARE OLD (✓) (RED IS YOUNGER THAN ORANGE AND OLDER THAN BLACK)	OLD is low saturation due to experiential correlations, UNSATURATED IS OLD - SATURATED IS YOUNG ✗ (no)	OLD is darkness due to experiential correlations, DARK IS OLD - BRIGHT IS YOUNG) ✓ (DARK IS OLD - BRIGHT IS YOUNG)

Table 5.12.: *Image Schemas and HSB Dependencies Predicted by Conceptual Metaphor Theory of Color in Detail, Empirical Results in Parentheses for Comparison (Continued)*

Image Schema	Hue-dependent	Saturation-dependent	Brightness-dependent
PAINFUL-NOT PAINFUL	red and black are associated with PAINFUL and green, blue, yellow are associated with NOT PAINFUL due to experiential correlations (RED, BLACK ARE PAINFUL - GREEN, BLUE. YELLOW ARE NOT PAINFUL) ✓ (RED IS MORE PAINFUL THAN CYAN, GREEN)	not painful is less saturated due to experiential correlations, SATURATED IS PAINFUL - UNSATURATED IS NOT PAINFUL ✓ (SATURATED IS PAINFUL - UNSATURATED IS NOT PAINFUL)	/ ✓ (no)
SMELLS GOOD-SMELLS BAD	/ ✗ (RED SMELLS BETTER THAN YELLOW, ORANGE)	/ ✓ (no)	SMELLS GOOD is brightness due to experiential correlations, BRIGHT SMELLS GOOD - DARK SMELLS BAD ✓ (BRIGHT SMELLS GOOD - DARK SMELLS BAD)
SMOOTH-ROUGH	SMOOTH is high surface reflection, e.g., WHITE IS SMOOTH ✗ (CYAN IS SMOOTHER THAN RED)	/ ✓ (no)	SMOOTH is high surface reflection, BRIGHT IS SMOOTH - DARK IS ROUGH ✓ (BRIGHT IS SMOOTH - DARK IS ROUGH)
STRONG-WEAK	STRONG is long wavelength hues and black due to cross-modal correspondences, BLACK IS STRONG ✓ (BLACK IS STRONGER THAN RED)	/ ✗ (SATURATED IS STRONG - UNSATURATED IS WEAK)	STRONG is darkness due to sexual dimorphism in skin pigmentation, DARK IS STRONG - BRIGHT IS WEAK ✓ (DARK IS STRONG - BRIGHT IS WEAK)

Table 5.12.: *Image Schemas and HSB Dependencies Predicted by Conceptual Metaphor Theory of Color in Detail, Empirical Results in Parentheses for Comparison (Continued)*

Image Schema	Hue-dependent	Saturation-dependent	Brightness-dependent
TASTES GOOD-TASTES BAD	hue of food often differs with sugar content, e.g., RED, ORANGE TASTE GOOD - GREEN, YELLOW, BROWN, BLACK, GRAY TASTE BAD ✓ (RED TASTES BETTER THAN BLUE, CYAN, VIOLET, GRAY)	/ ✓ (no)	TASTES GOOD is bright due to experiential correlations and cross-modal correspondences, BRIGHT TASTES GOOD - DARK TASTES BAD ✓ (BRIGHT TASTES GOOD - DARK TASTES BAD)
WARM-COLD	WARM is longer wavelength hues due to experiential correlations, e.g., RED IS WARM - BLUE IS COLD ✓ (RED IS WARMER THAN ALL OTHER HUES EXCEPT ORANGE)	/ ✓ (no)	/ ✗ (BRIGHT IS WARM - DARK IS COLD)

Note. /: no clear relationship found in literature. Results are marked with ✓ if the hypothesis was confirmed, and ✗if not confirmed. Results that only partly support a hypothesis are indicated by (✓).

cultures rate concepts on the dimensions of evaluation, potency and affect. Adams and Osgood report a study on 23 different cultures who rated 7 different colors (white, gray, black, red, yellow, green, blue) and found apparently universal trends in affective meanings of color (Adams & Osgood, 1973). Therefore, if subjects of different cultures also rate image schematic concepts with similar EPA-scores, the to be expected overlap between different cultures is high. However, as no EPA-profiles of colors and image schemas are available for Germans and Japanese separately, it cannot be predicted where variations are likely to occur.

Ecological Valence Theory on Human Color Preferences. If people with different cultural backgrounds have different color preferences, they are likely to recall associations to their preferred (or least preferred) objects, which can bias their image schema-color ratings. The favorite colors of the Japanese subjects are depicted on the right side in Fig. 5.3 together with the favorite colors of the German participants in Study 1. A Chi-square test of independence was calculated comparing the frequency of the color preferences of both cultures. The color preferences varied between German and Japanese participants, $\mathcal{X}^2(9) = 20.08$, $p = .018$, which is a typical result found in cross-cultural surveys (Ou et al., 2012). Because of this, variations in image-schematic evaluations are likely to occur.

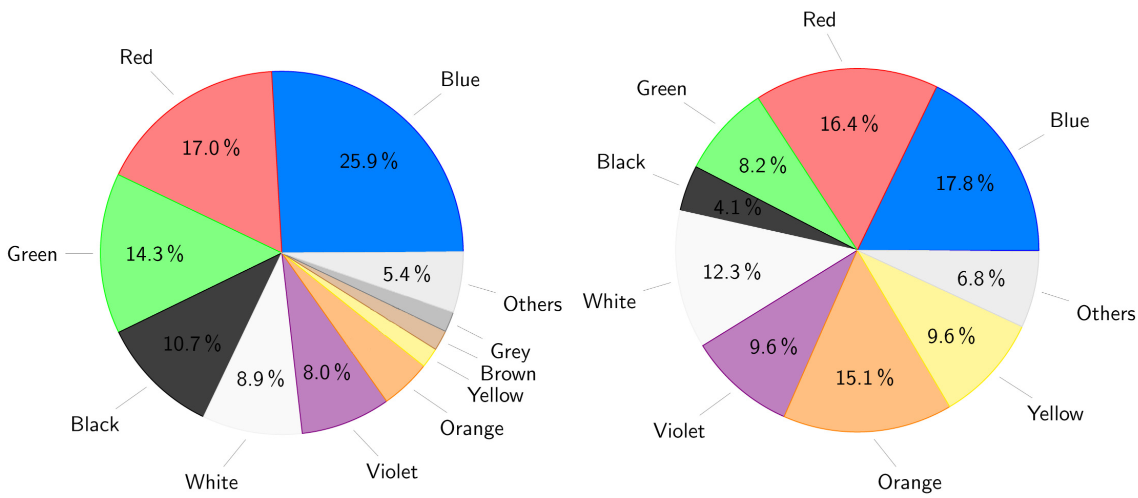

Figure 5.3.: Left: Favorite colors of German participants in Study 1 (percentages below 5% not labeled). Right: Favorite colors of Japanese participants in Study 2.

CMToC predicts that image-schematic color associations that are classified as *innate* and *sensorimotor* are universally acquired and thus shared across the human species. On the contrary, color associations that vary with the surrounding man-made environment (*culture*) are subject to intercultural and intracultural influences. Only PAINFUL-NOT PAINFUL is classified as a color association with cultural origin.

Indeed, PAINFUL-NOT PAINFUL was among the three image schemas with the lowest correlation coefficients. The observed difference in HARD-SOFT ratings indicates that subjects of both cultures probably encounter different colored materials more frequently. LOUD-SILENT would not be expected from this theoretical approach to be among the weakest correlations, as the association is assumed to be acquired through the common process of cross-modal correspondence between multimodal sensations of intensity. In German, people use synonyms of LOUD to describe eye-catching colors, e.g. *ein knalliges Rot* (a gaudy red). However, in Japanese, LOUD and related words are not used to define colors of specific properties. This means that although Japanese still have a quite similar pattern in ascribing image-schematic meaning to colors compared to Germans, the association is not surfaced in language.

5.2.4. Conclusion

The correlation coefficients between the 16 image-schematic ratings of German and Japanese participants all indicate a strong positive relationship with the exception of LOUD-SILENT. However, not only LOUD-SILENT but also the overlap in the image schemas HARD-SOFT and PAINFUL-NOT PAINFUL were considerably weaker than all other image schemas. Regarding the cultural stability of specific image-schematic color associations, theories in color psychology make only vague predictions. Color-in-context theory distinguishes innate and learned sources of color associations, but offers no guidance in determining the origin of a specific association in order to speculate about its cultural variability. Affective Meanings Systems predicts that if the EPA-profiles of image schemas and colors are similar between two different groups, the overlap will be high. As determined in their cross-cultural study involving 23 communities, Osgood and Adams found that the evaluations of most concepts remains very stable. This means that in general, a high overlap between the ratings of Germans and Japanese is to be expected, which is in line to what was found in Study 2. However, as the original publication does not comprise data for Germans and Japanese separately for the concepts of interest, image schemas and colors, it was not possible to derive more precise hypotheses. The Ecological Valence Theory on Human Color Preferences stresses the importance of the subjects' color preferences for the formation of color associations. Color associations are likely biased by recalling experiences with highly preferred or least preferred (colored) objects. As the color preferences between both cultural experimental groups differed, variations in image-schematic ratings were expected.

Finally, the theory developed within this work, CMToC, predicted a generally high overlap between Germans and Japanese in assigning image schemas to colors, as most associations originate from innate and sensorimotor experiences that are shared by both cultures. Moreover, one image schema-color association was expected to show a smaller overlap: PAINFUL-NOT PAINFUL. Indeed, PAINFUL-NOT was among the three image schemas with the lowest agreement. What was not initially predicted was the lower agreement for HARD-SOFT and LOUD-SILENT. However, after examining linguistic expressions involving color and the image schema, it was found for LOUD-SILENT that there are no cross-modal expressions that link both concepts, indicating a weaker link in Japanese subjects in general.

It is concluded that CMToC makes the most valid and precise predictions about the cultural stability and variation of semantic color associations as studied in this study paradigm. For those image schemas that originate from cultural experience (PAINFUL-NOT PAINFUL), the analysis of metaphoric speech can stimulate hypotheses about underlying relationships in a specific culture. Such a linguistic analysis can also shed light on image schemas that are less associated with color in a given culture (in case of Japanese LOUD-SILENT). With the literature compiled in this work and own skills to extrapolate from linguistic expressions to conceptual metaphors involving color, researchers and designers are well equipped to identify associations between colors and physical properties that are less prone to cultural influences. As Study 1 and 2 were the first to investigate the relationships between color and image schemas at large scale, a replication of the results in a different setting and with different participants is desirable. Two such conceptual replication attempts of the predictions of CMToC are described in the following sections.

5.3. Study 3: Confirmation of Color-Image Schema Associations with German Subjects

Studies 3 and 4 aim at a conceptual replication of the results of the previous two studies on the relationship between image schemas and the color attributes hue, saturation and brightness with different samples of German and Japanese subjects. The experimental paradigm is extended to include two-color samples to test color population stereotypes of physical properties. Population stereotypes describe ways in which a large proportion of a given population expects interface elements to function (Bergum & Bergum, 1981). Such population stereotypes provide practical rules that encourage user interface designs that match the mental models of the user and facilitate intuitive use (Hurtienne et al., 2009). Although much work on population stereotypes regarding the use of color has been done in the 1980s (Bergum & Bergum, 1981; Courtney, 1986), they are still an established procedure in HCI and are fundamental parts in today's ergonomic text books (Hurtienne et al., 2009). Often, population stereotypes have been investigated rather ad-hoc without a sound theoretical basis. CMToC on the contrary, enables designers to find new guidance for aligning color and physical properties grounded on a sound theoretical basis.

For this purpose, two online surveys were designed which oppose color patches that have high vs. low ratings on image schema dimensions. The hypothesis of these two studies is that when color pairs are more strongly associated with image schemas, participants will match them non-randomly to the image schema poles. The direction of the mapping can be predicted by the regression models of Study 1. If the colors are not or weakly associated with the image schema poles in question, they will be assigned by chance. The experimental setup, results and discussion of Study 3

(German subjects) and Study 4 (Japanese subjects) are described in the following.

5.3.1. Method

5.3.1.1. Participants

60 students from University of Würzburg, Germany, were recruited through an online recruiting system in exchange for course credit. Students with self-reported defective color vision did not participate. 8 data sets were excluded from data analysis because either the survey was not completed or participants made more than one error in three plates of the Ishihara Color Test (Ishihara, 1917). The remaining 52 participant's ages ranged from 18 to 24 years ($M = 20.8$, $SD = 1.7$). All participants (37 female, 15 male) had German as their mother tongue. Their favorite colors were blue (28.8%), red (25.0%), orange (3.8%), green (17.3%), violet (7.7%), yellow (1.9%), black (1.9%), gray (1.9%) and others (11.5%).

5.3.1.2. Procedure

Before working through the main part of the online survey, subjects had to complete three plates of the Ishihara Color Test (Ishihara, 1917) and a demographic questionnaire. In each trial, participants saw two color patches next to each other and, below them, two possibilities of how to assign image schematic poles. Their task was to choose which image schema assignment went best with the color composition in question (see Figure Fig. 5.4 for a sample screen and sec. A.3 for the instructions). The color patches were shown against gray background (HSB = n.d.,0,87). Participants made their choices by clicking on one of the two optional assignments with their preferred input device. The order of the color patches, image schema assignments and the left-right orientation of the image schema scales were randomized across participants. The survey took about 20 minutes to complete.

5.3.1.3. Material

Stimuli The online survey was created using EFS Survey. Shown color patches measured 400x400 pixel (length x width) and were displayed on the top center of the screen next to each other with a gap of 65 pixel in between. Below the color patches, two image-schematic attribute assignments were shown, matching one attribute either to the left or right color patch. For each of the 11 investigated image schemas[2] the labels were shown in 11pt Arial font.

[2] Note that the five image schemas newly added by Jörn Hurtienne (personal communication, March 12, 2013) were not investigated in this study: DIRTY-CLEAN, LOUD-SILENT, PAINFUL-NOT PAINFUL, SMELLS GOOD-SMELLS BAD, TASTES GOOD-TASTES BAD. However, their explorative character should encourage further research in order to gather empirical data to confirm them.

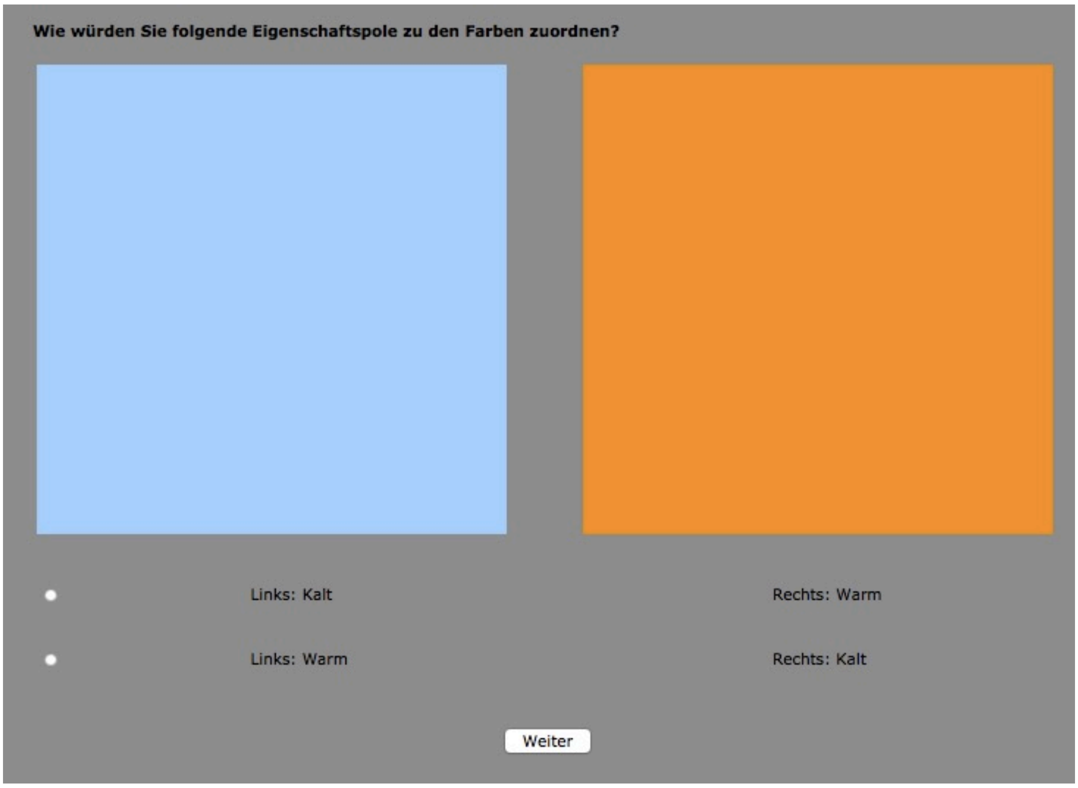

Figure 5.4.: Sample screen of Study 2. Translated from German: How would you assign the following dimensions to the depicted colors? Option 1: Left: Cold, Right: Warm. Option 2: Left: Warm, Right: Cold.

For each image schema pole, three colors with the highest ratings in Study 1 were chosen, as well as a pair of colors that scored neutral for the whole image schematic dimension, functioning as a control condition. Each of the three colors for an image-schematic pole was combined with the three colors of its opposite pole, resulting in nine test color pairs per image schema and a control pair. The selected colors are depicted in Tab. 5.13, together with their predicted image schema assignment based on the regression models from Study 1. If the z-score is negative, the color is associated with the unmarked image schema pole. Positive values are associated with the marked image schema pole.

Table 5.13.: *Color Patches, Related HSB Values and their Predicted z-Score on the Image Schema Dimension*

Image Schema	Color	H	S	B	Color Swatch	Predicted IS z-Score
BIG	black	n.d.	0	0		-0.742
	blue	212	90	25		-0.628
	red	0	100	75		-0.133
SMALL	violet	282	35	100		0.717
	green	130	35	90		0.542
	red	0	35	100		0.587
neutral	gray	n.d.	0	50		0.165
	cyan	191	75	70		0.072
BRIGHT	white	n.d.	0	100		-1.591
	yellow	56	35	100		-1.096
	blue	212	35	100		-0.705
DARK	black	n.d.	0	0		2.395
	blue	212	90	25		1.725
	brown	35	80	25		n/a*
neutral	green	130	100	50		0.334
	blue	212	75	100		-0.465
FAR	gray	n.d.	0	50		-0.580
	blue	212	35	100		-0.236
	blue	212	75	100		-0.036
NEAR	red	0	100	75		0.681
	orange	36	100	100		0.458
	yellow	56	100	95		0.324
neutral	green	130	35	90		0.079
	violet	282	35	100		0.118
FAST	yellow	56	100	95		-0.846
	red	0	100	75		-0.664
	cyan	181	75	70		-0.453
SLOW	brown	35	80	25		n/a*
	gray	n.d.	0	50		0.375
	green	130	75	25		0.055
neutral	yellow	56	35	100		-0.526
	green	130	35	90		-0.615

Table 5.13.: *Color Patches, Related HSB Values and their Predicted z-Score on the Image Schema Dimension (Continued)*

Image Schema	Color	H	S	B	Color Swatch	Predicted IS z-Score
FULL	red	0	100	25		-1.583
	blue	212	90	25		-1.230
	black	n.d.	0	0		-1.054
EMPTY	white	n.d.	0	100		0.192
	blue	212	35	100		0.050
	orange	36	35	100		0.078
neutral	cyan	181	75	70		-0.782
	red	0	35	100		-0.193
HARD	black	n.d.	0	0		-1.495
	brown	35	80	25		n/a*
	violet	282	100	25		-0.616
SOFT	red	0	35	100		0.195
	violet	282	35	100		0.224
	orange	36	35	100		0.185
neutral	yellow	56	100	95		-0.367
	green	130	100	50		-0.706
HEAVY	black	n.d.	0	0		-1.155
	red	0	100	25		-1.049
	blue	212	90	25		-0.855
LIGHT	white	n.d.	0	100		1.798
	red	0	35	100		1.141
	blue	212	35	100		1.275
neutral	green	130	100	50		-0.081
	brown	35	80	25		n/a*
OLD	brown	35	80	25		n/a*
	brown	35	80	50		n/a*
	gray	n.d.	0	50		-0.534
YOUNG	yellow	56	100	95		1.008
	red	0	35	100		1.119
	blue	212	35	100		1.165
neutral	cyan	181	100	35		0.216
	yellow	50	100	85		0.798

Table 5.13.: *Color Patches, Related HSB Values and their Predicted z-Score on the Image Schema Dimension (Continued)*

Image Schema	Hue	H	S	B	Color Swatch	Predicted IS Score
SMOOTH	white	n.d.	0	100		0.375
	blue	212	35	100		0.342
	yellow	56	100	95		-0.352
ROUGH	gray	n.d.	0	50		1.168
	brown	35	80	25		n/a*
	brown	35	80	50		n/a*
neutral	cyan	181	100	35		-1.002
	orange	30	100	75		-0.392
STRONG	black	n.d.	0	0		-1.321
	red	0	100	75		-0.612
	red	0	100	25		-0.912
WEAK	orange	36	35	100		0.667
	red	0	35	100		0.513
	violet	282	35	100		0.805
neutral	white	n.d.	0	100		0.262
	gray	n.d.	0	50		0.755
WARM	orange	36	100	100		-1.076
	red	0	100	25		-1.203
	red	0	100	75		-1.353
COLD	blue	212	35	100		0.352
	gray	n.d.	0	50		0.492
	white	n.d.	0	100		0.182
neutral	green	130	100	50		-0.317
	violet	282	35	100		-0.445

**Note.* No data available for brown hues since they were excluded from the regression model because of the interdependence between HSB for this color. Brown is still included in the stimulus set of Study 2 because of strong associations with certain image schemas.

Auxiliary data In the preliminary questionnaire, the plate numbers 1 (showing *12*), 7 (showing *74*), 8 (showing *6*) of the Ishihara Color Test (Ishihara, 1917) were used to test for defective color vision. The subsequent demographic questionnaire had to be completed with age, gender, current occupation, highest level of education, ethnicity, mother tongue and favorite color.

5.3.1.4. Experimental Design and Data Analysis

First, Cohen's kappa values as an index of mapping strength ($\varkappa$) were calculated for each condition and image schema (image schema-related color pair (called 'test' condition from now) vs. image schema-non-related color pair). The $\varkappa$-ratio represents the number of choices in line with the predicted image-schematic color mapping while considering agreement by chance (50%) (Chan, Shum, Law, & Hui, 2003). $\varkappa$-values range from -1 (perfect disagreement with the mapping) to 1 (perfect agreement with the mapping). An agreement value above .40 can be interpreted as 'moderate', above .60 as 'substantial', and above .80 as 'almost perfect' (Landis & Koch, 1977). Subsequently, the within-subjects design was subjected to multiple *t*-tests with the independent variable color pair (image schema-related (test) vs. neutral) and kappa values as the dependent variable.

5.3.2. Results

The non-averaged $\varkappa$-vales of every color pair can be found in Tab. 5.14. The averaged $\varkappa$-values and percentages for the test color pairs as well as the neutral control conditions are summarized in Tab. 5.15 .

Table 5.14.: *%- and κ-Values for Test and Control Color Pairs Together with the Predicted Absolute Distance on the Image Schema Dimension*

Image Schema	H^1	S^1	B^1	H^2	S^2	B^2	Color Swatches$^{1\text{-}2}$	%	ϰ	z-Score Distance
BIG-	n.d.	0	0	282	35	100		67	0.34	1.459
SMALL	n.d.	0	0	130	35	90		67	0.34	1.284
	n.d.	0	0	0	35	100		77	0.54	1.329
	212	90	25	282	35	100		69	0.38	1.345
	212	90	25	130	35	90		71	0.42	1.170
	212	90	25	0	35	100		71	0.42	1.215
	0	100	75	282	35	100		67	0.34	0.850
	0	100	75	130	35	90		67	0.35	0.675
	0	100	75	0	35	100		75	0.50	0.720
(neutral)	n.d.	0	50	191	75	70		46	-0.08	0.093
BRIGHT-	n.d.	0	100	n.d.	0	0		98	0.96	3.986
DARK	n.d.	0	100	212	90	25		100	1.00	3.316
	n.d.	0	100	35	80	25		98	0.96	n/a*
	56	35	100	n.d.	0	0		96	0.92	3.491
	56	35	100	212	90	25		100	1.00	2.821
	56	35	100	35	80	25		98	0.96	n/a*
	212	35	100	n.d.	0	0		100	1.00	3.100
	212	35	100	212	90	25		98	0.96	2.430
	212	35	100	35	80	25		100	1.00	n/a*
(neutral)	130	100	50	212	75	100		42	-0.15	0.799
FAR-	n.d.	0	50	0	100	75		98	0.96	1.261
NEAR	n.d.	0	50	36	100	100		90	0.81	1.038
	n.d.	0	50	56	100	95		88	0.77	0.904
	212	35	100	0	100	75		92	0.85	0.917
	212	35	100	36	100	100		94	0.88	0.694
	212	35	100	56	100	95		92	0.85	0.560
	212	75	100	0	100	75		96	0.92	0.717
	212	75	100	36	100	100		77	0.54	0.494
	212	75	100	56	100	95		83	0.65	0.360
(neutral)	130	35	90	282	35	100		48	-0.04	0.039

Table 5.14.: *%- and κ-Values for Test and Control Color Pairs Together with the Predicted Absolute Distance on the Image Schema Dimension*

Image Schema	H[1]	S[1]	B[1]	H[2]	S[2]	B[2]	Color Swatches[1-2]	%	κ	z-Score Distance
FAST-	56	100	95	35	80	25		90	0.81	n/a*
SLOW	56	100	95	n.d.	0	50		98	0.96	1.211
	56	100	95	130	75	25		98	0.96	0.901
	0	100	75	35	80	25		94	0.88	n/a*
	0	100	75	n.d.	0	50		90	0.81	1.039
	0	100	75	130	75	25		82	0.65	0.719
	181	75	70	35	80	25		90	0.81	n/a*
	181	75	70	n.d.	0	50		90	0.81	0.828
	181	75	70	130	75	25		90	0.81	0.508
(neutral)	56	35	100	130	35	90		65	0.31	0.089
FULL-	0	100	25	n.d.	0	100		96	0.92	1.775
EMPTY	0	100	25	212	35	100		98	0.96	1.633
	0	100	25	36	35	100		90	0.81	1.661
	212	90	25	n.d.	0	100		94	0.88	1.422
	212	90	25	212	35	100		94	0.88	1.280
	212	90	25	36	35	100		90	0.81	1.308
	n.d.	0	0	n.d.	0	100		94	0.88	1.246
	n.d.	0	0	212	35	100		88	0.77	1.104
	n.d.	0	0	36	35	100		88	0.77	1.132
(neutral)	181	75	70	0	35	100		27	-0.46	0.589
HARD-	n.d.	0	0	0	35	100		100	1.00	1.690
SOFT	n.d.	0	0	282	35	100		92	0.85	1.719
	n.d.	0	0	36	35	100		96	0.92	1.680
	35	80	25	0	35	100		94	0.88	n/a*
	35	80	25	282	35	100		83	0.65	n/a*
	35	80	25	36	35	100		92	0.85	n/a*
	282	100	25	0	35	100		83	0.65	0.840
	282	100	25	282	35	100		83	0.65	0.840
	282	100	25	36	35	100		85	0.69	0.801
(neutral)	56	100	95	130	100	50		40	-0.19	0.339

Table 5.14.: *%- and κ-Values for Test and Control Color Pairs Together with the Predicted Absolute Distance on the Image Schema Dimension (Continued)*

Image Schema	H^1	S^1	B^1	H^2	S^2	B^2	Color Swatches$_{1\text{-}2}$	%	κ	z-Score Distance
HEAVY-	n.d.	0	0	n.d.	0	100		62	0.23	2.953
LIGHT	n.d.	0	0	0	35	100		98	0.96	2.296
	n.d.	0	0	212	35	100		96	0.92	2.430
	0	100	25	n.d.	0	100		98	0.96	2.847
	0	100	25	0	35	100		100	1.00	2.190
	0	100	25	212	35	100		98	0.96	2.324
	212	90	25	n.d.	0	100		98	0.96	2.653
	212	90	25	0	35	100		100	1.00	1.996
	212	90	25	212	35	100		100	1.00	2.130
(neutral)	130	100	50	35	80	25		19	-0.61	n/a*
OLD-	35	80	25	56	100	95		96	0.92	n/a*
YOUNG	35	80	25	0	35	100		96	0.92	n/a*
	35	80	25	212	35	100		94	0.88	n/a*
	35	80	50	56	100	95		96	0.92	n/a*
	35	80	50	0	35	100		94	0.88	n/a*
	35	80	50	212	35	100		96	0.92	n/a*
	n.d.	0	50	56	100	95		94	0.88	1.542
	n.d.	0	50	0	35	100		94	0.88	1.653
	n.d.	0	50	212	35	100		96	0.92	1.699
(neutral)	181	100	35	50	100	85		71	0.42	0.582
SMOOTH-	n.d.	0	100	n.d.	0	50		87	0.73	0.793
ROUGH	n.d.	0	100	35	80	25		92	0.85	n/a*
	n.d.	0	100	35	80	50		96	0.92	n/a*
	212	35	100	n.d.	0	50		90	0.81	0.826
	212	35	100	35	80	25		96	0.92	n/a*
	212	35	100	35	80	50		96	0.92	n/a*
	56	100	95	n.d.	0	50		92	0.85	0.816
	56	100	95	35	80	25		88	0.77	n/a*
	56	100	95	35	80	50		92	0.85	n/a*
(neutral)	181	100	35	30	100	75		40	-0.19	0.610

Table 5.14.: *%- and κ-Values for Test and Control Color Pairs Together with the Predicted Absolute Distance on the Image Schema Dimension (Continued)*

Image Schema	H[1]	S[1]	B[1]	H[2]	S[2]	B[2]	Color Swatches[1-2]	%	κ	z-Score Distance
STRONG-	n.d.	0	0	36	35	100		92	0.85	1.988
WEAK	n.d.	0	0	0	35	100		94	0.88	1.834
	n.d.	0	0	282	35	100		94	0.88	2.126
	0	100	75	36	35	100		94	0.88	0.667
	0	100	75	0	35	100		98	0.96	1.125
	0	100	75	282	35	100		92	0.85	1.417
	0	100	25	36	35	100		96	0.92	1.579
	0	100	25	0	35	100		94	0.88	1.425
	0	100	25	282	35	100		90	0.81	1.717
(neutral)	n.d.	0	100	n.d.	0	50		25	-0.5	0.493
WARM-	36	100	100	212	35	100		1	1.00	1.428
COLD	36	100	100	n.d.	0	50		1	1.00	1.566
	36	100	100	n.d.	0	100		1	1.00	1.258
	0	100	25	212	35	100		92	0.85	1.555
	0	100	25	n.d.	0	50		94	0.88	1.695
	0	100	25	n.d.	0	100		94	0.88	1.385
	0	100	75	212	35	100		98	0.96	1.705
	0	100	75	n.d.	0	50		1	1.00	1.845
	0	100	75	n.d.	0	100		96	0.92	1.535
(neutral)	130	100	50	282	35	100		52	0.04	0.128

**Note.* Distance could not be calculated since no data was available for brown hues.

Table 5.15.: *Averaged %- and κ-Values for Test and Control Color Pairs*

Image Schema	Condition	%	κ
BIG-SMALL	Test	70	0.40
	Control	46	-0.08
BRIGHT-DARK	Test	99	0.86
	Control	42	-0.15
FAR-NEAR	Test	90	0.80
	Control	48	-0.04
FAST-SLOW	Test	91	0.83
	Control	65	0.31
FULL-EMPTY	Test	92	0.85
	Control	27	-0.46
HARD-SOFT	Test	90	0.79
	Control	40	-0.19
HEAVY-LIGHT	Test	94	0.89
	Control	19	-0.61
OLD-YOUNG	Test	95	0.90
	Control	71	0.42
SMOOTH-ROUGH	Test	92	0.85
	Control	40	-0.19
STRONG-WEAK	Test	94	0.88
	Control	25	-0.5
WARM-COLD	Test	97	0.94
	Control	52	0.04
Total	**Test**	**91**	**0.82**
	Control	**43**	**-0.13**

Except for BIG-SMALL, the percentages of participants' choices in favor of the predicted image-schematic color mappings are all 90% or above and κ-values are equal or larger than 0.80. Thus, BIG-SMALL mappings were identified with 'moderate' agreement, whereas all other test color pairs achieved 'almost perfect' agreement with the related image schema. Using two one-sample *t*-tests, the averaged κ-values for all test conditions as well as control conditions were compared to agreement by chance ($\varkappa = 0.00$). Whereas the test condition (M = 0.82, SD = 0.15) differed significantly from random selection, $t(10) = 18.69$, $p < .001$, $d = 5.47$, the control condition (M = -0.13, SD = 0.32) did not, $t(10) = -1.37$, $p = .201$.

After averaging percentage scores and ϰ-values of both conditions across all image schemas, the result was subjected to a paired-samples t-test. There was a significant difference between both the percentage values of test (M = 91.27, SD = 7.60) and control condition (M = 43.18, SD = 15.99), $t(10) = 8.86$, $p < .001$, $d = 3.84$, as well as the ϰ-values of test (M = 0.82, SD = 0.15) and control condition (M = -0.13, SD = 0.32), $t(10) = 8.83$, $p < .001$, $d = 3.80$.

Next, the relationship between the predicted z-scores of single colors on an image schema dimension and the choices in assigning selected color pairs to image schema poles was inspected. Therefore, the scores of each of the color patches from the regression analysis of Study 1 and the kappa-values of test and control condition (see Tab. 5.14) were subjected to a correlation analysis. For this purpose, the absolute z-score distance between the two color samples of each color pair was calculated (Tab. 5.14). It is expected that the greater the distance between the two colors on an image schema dimension and therefore the greater the difference from 0, the more unambiguous the choices of the participants, i.e., above chance level. The results support the hypothesis and show a strong positive relationship, $r(21) = .663$, $p = .001$.

5.3.3. Discussion

Overall, different participants reliably assigned the image schema dimensions to the color pairs as predicted by the regression models of Study 1. Although BIG-SMALL received only 'moderate' agreement regarding the matched test color pairs, the other image schemas performed 'almost perfect' and thus qualify as population stereotypes. In contrast, the color pairs in the control condition were assigned randomly, suggesting that the association with the specific image schema are weak or not existent. An explanation regarding the poor performance of the BIG-SMALL image schema might be the foreground-background contrast ratio, as already discussed in sec. 5.1.4. While CMToC would predict, given a high foreground-background contrast ratio, that brighter colors are judged larger, the gray background used in all surveys reduced the foreground-background contrast ratio for bright colors so that participants' judgements were reversed. Given that, the gray background might have ambiguated color effects on size judgements. Moreover, since in Study 3 two equally sized color patches were presented next to each other, participants could directly compare the size of both, which could have also weakened the overall impact of color.

5.4. Study 4: Confirmation of Color-Image Schema Associations with Japanese Subjects

Study 4 is a confirmative replication attempt of Study 3 with Japanese subjects. In the following, differences in the method compared to Study 3 with German subjects are highlighted, results are presented and discussed.

5.4.1. Method

5.4.1.1. Participants

46 students from the campus of Nihon University, Japan, volunteered to participate in the study in exchange for course credit. 9 data sets were excluded from data analysis because either the survey was not completed or participants made more than one error in three plates of the Ishihara Color Test (Ishihara, 1917). The remaining 37 participants (24 female, 13 male) were between 19-21 years old (M = 19.5, SD = 0.56), had Japanese as their mother tongue, and had no reported defective color vision. Their favorite colors were blue (27.0%), red (18.9%), green (10.8%), white (8.1%), violet (5.4%), orange (2.7%), yellow (2.7%), and others (21.6%).

5.4.1.2. Procedure

The online survey and experimental procedure were identical to Study 3 (sec. 5.3.1 for details). After the participants completed the three plates of the Ishihara Color Test (Ishihara, 1917) and the demographic questionnaire, they had to work through the main experiment. Again, the experimental trials were preceded by a practice trial that explained the task in detail (see sec. A.4 for the instruction). In each trial, participants had to match two colored patches with a pair of image schemas. Completing the 110 trials took about 20 minutes.

5.4.1.3. Material

Stimuli The stimulus material was similar to Study 3 (see sec. 5.3.1). The translated names of the 11 image schema were taken from Study 1 (see Tab. A.5). For each image schema, three colors with the highest ratings of Study 2 were chosen (test colors) and a control color pair with a rating around the neutral midpoint of the scale. Therefore, by combining the 3x3 test colors with each other and the control color pair, 10 color pairs were shown for each image schema in total. The color patches selected for this study are shown in Tab. 5.16, including their predicted image schema assignment based on the regression models from Study 1 (German subjects) and Study 2 (Japanese subjects). Overall, 38% of the chosen color patches differed from

Study 3 (22% concerned test colors and 14% concerned control colors). Note that if the z-score is negative, the color is associated with the unmarked image schema pole. Positive values are associated with the marked image schema pole.

Table 5.16.: *Color Patches, Related HSB Values and their Predicted z-Score on the Image Schema Dimension*

Image Schema	Color	H	S	B	Color Swatch	Predicted IS z-Score Japanese Data	Predicted IS z-Score German Data
BIG	black	n.d.	0	0		-1.030	-0.742
	red	0	100	25		-0.645	-1.583
	red	0	100	75		-0.295	-0.133
SMALL	yellow	56	35	100		0.513	0.634
	green	130	35	90		0.426	0.542
	red	0	35	100		0.270	0.587
neutral	gray	n.d.	0	50		0.259	0.165
	cyan	181	75	70		0.087	0.072
BRIGHT	white	n.d.	0	100		-1.359	-1.591
	yellow	56	35	100		-0.855	-1.096
	cyan	181	35	100		-0.949	-0.990
DARK	black	n.d.	0	0		1.954	2.395
	blue	212	90	25		1.537	1.725
	brown	35	80	25		n/a*	n/a*
neutral	green	130	100	50		0.242	0.334
	blue	212	75	100		-0.578	-0.465
FAR	gray	n.d.	0	50		-0.551	-0.580
	cyan	181	100	35		-0.192	-0.108
	blue	212	75	100		0.035	-0.036
NEAR	red	0	100	75		0.487	0.681
	red	0	35	100		0.427	0.456
	yellow	56	100	95		0.490	0.324
neutral	brown	35	50	75		n/a*	n/a*
	cyan	181	35	100		0.068	-0.173
FAST	yellow	56	100	95		-0.870	-0.846
	red	0	100	75		-0.596	-0.664
	cyan	181	35	100		-0.530	-0.644
SLOW	brown	35	80	25		n/a*	n/a*
	brown	35	80	50		n/a*	n/a*
	green	130	75	25		0.279	0.055
neutral	yellow	56	35	100		-0.700	-0.526
	green	130	35	90		-0.731	-0.615

Table 5.16.: *Color Patches, Related HSB Values and their Predicted z-Score on the Image Schema Dimension (Continued)*

Image Schema	Color	H	S	B	Color Swatch	Predicted IS z-Score Japanese Data	Predicted IS z-Score German Data
FULL	red	0	100	25		-1.410	-1.583
	red	0	100	75		-0.910	-1.353
	orange	36	100	100		-0.472	-0.637
EMPTY	white	n.d.	0	100		-0.320	0.192
	blue	212	35	100		0.297	0.050
	cyan	181	35	100		0.401	-0.072
neutral	orange	36	35	100		-0.017	0.078
	brown	35	50	75		n/a*	n/a*
HARD	black	n.d.	0	0		-1.554	-1.495
	brown	35	80	25		n/a*	n/a*
	violet	282	100	25		-0.958	-0.616
SOFT	red	0	35	100		0.331	0.195
	violet	282	35	100		0.342	0.224
	orange	36	35	100		0.526	0.185
neutral	brown	35	80	50		n/a*	n/a*
	green	130	75	25		-0.480	0.055
HEAVY	black	n.d.	0	0		-0.435	-1.155
	red	0	100	25		-0.476	-1.049
	blue	212	90	25		-0.540	-0.855
LIGHT	white	n.d.	0	100		2.382	1.798
	red	0	35	100		2.356	1.141
	cyan	181	35	100		1.587	1.512
neutral	orange	36	100	100		1.178	0.458
	brown	35	50	75		n/a*	n/a*
OLD	brown	35	80	25		n/a*	n/a*
	brown	35	80	50		n/a*	n/a*
	red	0	100	25		-0.409	-1.049
YOUNG	yellow	56	100	95		1.040	1.008
	red	0	35	100		1.346	1.119
	cyan	181	35	100		1.328	1.321
neutral	green	130	100	50		0.370	0.334
	violet	282	35	100		0.930	0.224

Table 5.16.: *Color Patches, Related HSB Values and their Predicted z-Score on the Image Schema Dimension (Continued)*

Image Schema	Color	H	S	B	Color Swathe	Predicted IS z-Score Japanese Data	Predicted IS z-Score German Data
SMOOTH	white	n.d.	0	100		-1.350	0.375
	blue	212	35	100		-0.929	0.342
	cyan	181	35	100		-1.159	0.363
ROUGH	green	130	75	25		0.056	0.055
	brown	35	80	25		n/a*	n/a*
	brown	35	80	50		n/a*	n/a*
neutral	yellow	50	100	85		-0.398	-0.412
	orange	36	100	100		-0.578	-0.242
STRONG	black	n.d.	0	0		-1.628	-1.321
	red	0	100	75		-1.131	-0.612
	red	0	100	25		-1.481	-0.912
WEAK	orange	36	35	100		0.043	0.667
	red	0	35	100		-0.176	0.513
	cyan	181	35	100		0.455	0.836
neutral	blue	212	75	100		-0.415	-0.036
	brown	35	80	50		n/a*	n/a*
WARM	orange	36	100	100		-1.263	-1.076
	red	0	100	75		-1.307	-1.203
	orange	30	100	75		-1.088	-0.392
COLD	blue	212	35	100		0.424	0.352
	gray	n.d.	0	50		0.399	0.492
	blue	212	75	100		0.344	-0.036
neutral	green	130	100	50		-0.276	-0.317
	violet	282	35	100		-0.386	-0.445

**Note.* No data available for brown hues since they were excluded from the regression model because of the interdependence between HSB for this color. Brown is still included in the stimulus set of Study 2 because of strong associations with certain image schemas.

Auxiliary data The preliminary questionnaire contained the same plates of the Ishihara Color Test (Ishihara, 1917) as in Study 3. In the demographic questionnaire, participants had to state age, gender, current occupation, highest level of education, ethnicity, mother tongue and favorite color.

5.4.1.4. Experimental Design and Data Analysis

The within-subjects experimental design and data analysis were identical to Study 3. Percentages and $\varkappa$-values were computed for the participants' choices. Percentages around 50% and $\varkappa$-values around 0 indicate random choices, whereas values below/above 0 indicate a specific association of the presented color pair with the image schema domain. It is expected that the test color pairs are reliably associated with the image schema domain in focus of the task, whereas the participants' choices in the control color pair condition will be at chance level.

5.4.2. Results

The non-averaged data is depicted in Tab. 5.17 and Tab. 5.18 summarizes mean $\varkappa$-values and percentages for both test color pairs as well as the neutral control conditions.

Table 5.17.: *%- and κ-Values for Test and Control Color Pairs Together with the Predicted Absolute z-Score Distance on the Image Schema Dimension from Japanese (D_J) and German Data (D_G)*

Image Schema	H[1]	S[1]	B[1]	H[2]	S[2]	B[2]	Color Swatches[1-2]	%	κ	D_J	D_G
BIG-	n.d.	0	0	56	35	100		57	0.14	1.543	1.376
SMALL	n.d.	0	0	130	35	90		57	0.14	1.456	1.284
	n.d.	0	0	0	35	100		54	0.08	1.300	1.329
	0	100	25	56	35	100		62	0.24	1.158	2.217
	0	100	25	130	35	90		57	0.14	1.071	2.125
	0	100	25	0	35	100		65	0.30	0.915	2.170
	0	100	75	56	35	100		70	0.41	0.808	0.767
	0	100	75	130	35	90		68	0.35	0.721	0.675
	0	100	75	0	35	100		70	0.41	0.565	0.720
(neutral)	n.d.	0	50	181	75	70		24	-0.51	0.172	0.093
BRIGHT-	n.d.	0	100	n.d.	0	0		97	0.95	3.313	3.986
DARK	n.d.	0	100	212	90	25		100	1.00	2.896	3.316
	n.d.	0	100	35	80	25		100	1.00	n/a*	n/a*
	56	35	100	n.d.	0	0		97	0.95	2.809	3.491
	56	35	100	212	90	25		100	1.00	2.392	2.821
	56	35	100	35	80	25		97	0.95	n/a*	n/a*
	181	35	100	n.d.	0	0		100	1.00	2.903	3.385
	181	35	100	212	90	25		100	1.00	2.486	2.715
	181	35	100	35	80	25		100	1.00	n/a*	n/a*
(neutral)	130	100	50	212	75	100		30	-0.41	0.820	0.799
FAR-	n.d.	0	50	0	100	75		92	0.84	1,038	1.261
NEAR	n.d.	0	50	0	35	100		81	0.62	0,978	1.036
	n.d.	0	50	56	100	95		100	1.00	1,041	0.904
	181	100	35	0	100	75		89	0.78	0,679	0.789
	181	100	35	0	35	100		70	0.41	0,619	0.564
	181	100	35	56	100	95		86	0.73	0,682	0.432
	212	75	100	0	100	75		78	0.57	0,452	0.717
	212	75	100	0	35	100		59	0.19	0,392	0.492
	212	75	100	56	100	95		92	0.84	0,455	0.360
(neutral)	35	50	75	181	35	100		57	0.14	n/a*	n/a*

Table 5.17.: *%- and κ-Values for Test and Control Color Pairs Together with the Predicted Absolute z-Score Distance on the Image Schema Dimension from Japanese (D_J) and German Data (D_G)*

Image Schema	H[1]	S[1]	B[1]	H[2]	S[2]	B[2]	Color Swatches[1-2]	%	κ	D_J	D_G
FAST-	56	100	95	35	80	25		97	0.95	n/a*	n/a*
SLOW	56	100	95	35	80	50		95	0.89	n/a*	n/a*
	56	100	95	130	75	25		92	0.84	1.149	0.901
	0	100	75	35	80	25		91	0.84	n/a*	n/a*
	0	100	75	35	80	50		92	0.84	n/a*	n/a*
	0	100	75	130	75	25		92	0.84	0.875	0.719
	181	35	100	35	80	25		95	0.89	n/a*	n/a*
	181	35	100	35	80	50		68	0.35	n/a*	n/a*
	181	35	100	130	75	25		92	0.84	0.809	0.699
(neutral)	56	35	100	130	35	90		51	0.03	-0.031	0.089
FULL-	0	100	25	n.d.	0	100		95	0.89	1.090	1.775
EMPTY	0	100	25	212	35	100		92	0.84	1.707	1.633
	0	100	25	181	35	100		92	0.84	1.811	1.655
	0	100	75	n.d.	0	100		97	0.95	0.590	1.545
	0	100	75	212	35	100		97	0.95	1.207	1.403
	0	100	75	181	35	100		95	0.89	1.311	1.425
	36	100	100	n.d.	0	100		92	0.84	0.152	0.829
	36	100	100	212	35	100		97	0.95	0.769	0.687
	36	100	100	181	35	100		95	0.89	0.873	0.709
(neutral)	36	35	100	35	50	75		81	0.62	n/a*	n/a*
HARD-	n.d.	0	0	0	35	100		100	1.00	1.885	1.690
SOFT	n.d.	0	0	282	35	100		97	0.95	1.896	1.719
	n.d.	0	0	36	100	100		100	1.00	2.080	1.680
	35	80	25	0	35	100		97	0.95	n/a*	n/a*
	35	80	25	282	35	100		100	1.00	n/a*	n/a*
	35	80	25	36	100	100		100	1.00	n/a*	n/a*
	282	100	25	0	35	100		100	1.00	1.289	0.811
	282	100	25	282	35	100		97	0.95	1.300	0.840
	282	100	25	36	100	100		100	1.00	1.484	0.801
(neutral)	35	80	50	130	75	25		65	0.30	n/a*	n/a*

Table 5.17.: *%- and κ-Values for Test and Control Color Pairs Together with the Predicted Absolute z-Score Distance on the Image Schema Dimension from Japanese (D_J) and German Data (D_G) (Continued)*

Image Schema	H[1]	S[1]	B[1]	H[2]	S[2]	B[2]	Color Swatches[1-2]	%	ϰ	D_J	D_G
HEAVY-	n.d.	0	0	n.d.	0	100		97	0.95	2.817	2.953
LIGHT	n.d.	0	0	0	35	100		100	1.00	2.791	2.296
	n.d.	0	0	181	35	100		100	1.00	2.022	2.667
	0	100	25	n.d.	0	100		100	1.00	2.858	2.847
	0	100	25	0	35	100		100	1.00	2.832	2.190
	0	100	25	181	35	100		100	1.00	2.063	2.561
	212	90	25	n.d.	0	100		100	1.00	2.922	2.653
	212	90	25	0	35	100		100	1.00	2.896	1.996
	212	90	25	181	35	100		100	1.00	2.127	2.367
(neutral)	36	100	100	35	50	75		35	-0.30	n/a*	n/a*
OLD-	35	80	25	56	100	95		100	1.00	n/a*	n/a*
YOUNG	35	80	25	0	35	100		100	1.00	n/a*	n/a*
	35	80	25	181	35	100		97	0.95	n/a*	n/a*
	35	80	50	56	100	95		97	0.95	n/a*	n/a*
	35	80	50	0	35	100		97	0.95	n/a*	n/a*
	35	80	50	181	35	100		97	0.95	n/a*	n/a*
	0	100	25	56	100	95		100	1.00	1.449	2.057
	0	100	25	0	35	100		97	0.95	1.755	2.168
	0	100	25	181	35	100		95	0.89	1.737	2.370
(neutral)	130	100	50	282	35	100		54	0.08	0.560	0.110
SMOOTH-	n.d.	0	100	130	75	25		95	0.89	1.406	0.320
ROUGH	n.d.	0	100	35	80	25		89	0.78	n/a*	n/a*
	n.d.	0	100	35	80	50		83	0.68	n/a*	n/a*
	212	35	100	130	75	25		86	0.73	0.985	0.287
	212	35	100	35	80	25		89	0.78	n/a*	n/a*
	212	35	100	35	80	50		89	0.78	n/a*	n/a*
	181	35	100	130	75	25		92	0.84	1.215	0.308
	181	35	100	35	80	25		95	0.89	n/a*	n/a*
	181	35	100	35	80	50		92	0.84	n/a*	n/a*
(neutral)	50	100	85	36	100	100		49	-0.03	0.180	0.170

Table 5.17.: *%- and κ-Values for Test and Control Color Pairs Together with the Predicted Absolute z-Score Distance on the Image Schema Dimension from Japanese (D_J) and German Data (D_G) (Continued)*

Image Schema	H¹	S¹	B¹	H²	S²	B²	Color Swatches₁₋₂	%	κ	D_J	D_G
STRONG-	n.d.	0	0	36	100	100		97	0.95	1.671	1.988
WEAK	n.d.	0	0	0	35	100		97	0.95	1.452	1.834
	n.d.	0	0	181	35	100		97	0.95	2.083	2.157
	0	100	75	36	100	100		100	1.00	1.174	1.279
	0	100	75	0	35	100		97	0.95	0.955	1.125
	0	100	75	181	35	100		100	1.00	1.586	1.448
	0	100	25	36	100	100		100	1.00	1.524	1.579
	0	100	25	0	35	100		97	0.95	1.305	1.425
	0	100	25	181	35	100		89	0.78	1.936	1.748
(neutral)	212	75	100	35	80	50		41	-0.19	n/a*	n/a*
WARM-	36	100	100	212	35	100		100	1.00	1.687	1.428
COLD	36	100	100	n.d.	0	50		97	0.95	1.662	1.568
	36	100	100	212	75	100		97	0.95	1.607	1.040
	0	100	75	212	35	100		100	1.00	1.731	1.555
	0	100	75	n.d.	0	50		100	1.00	1.706	1.695
	0	100	75	212	75	100		97	0.95	1.651	1.167
	30	100	75	212	35	100		100	1.00	1.512	0.744
	30	100	75	n.d.	0	50		97	0.95	1.487	0.884
	30	100	75	212	75	100		95	0.89	1.432	0.356
(neutral)	130	100	50	282	35	100		43	-0.14	0.110	0.128

**Note.* Distance could not be calculated since no data was available for brown hues.

Table 5.18.: *Averaged %- and κ-Values for Test and Control Color Pairs*

Image Schema	Condition	%	ϰ
BIG-SMALL	Test	62	0.25
	Control	24	-0.51
BRIGHT-DARK	Test	99	0.98
	Control	30	-0.40
FAR-NEAR	Test	83	0.66
	Control	57	0.14
FAST-SLOW	Test	90	0.81
	Control	51	0.02
FULL-EMPTY	Test	95	0.89
	Control	81	0.62
HARD-SOFT	Test	99	0.98
	Control	65	0.30
HEAVY-LIGHT	Test	100	0.99
	Control	35	-0.30
OLD-YOUNG	Test	98	0.96
	Control	54	0.08
SMOOTH-ROUGH	Test	90	0.80
	Control	49	-0.02
STRONG-WEAK	Test	97	0.94
	Control	41	-0.18
WARM-COLD	Test	98	0.96
	Control	43	-0.14
Total	**Test**	**92**	**0.84**
	Control	**48**	**-0.04**

Similar to the results of Study 3, all image schemas except BIG-SMALL achieve percentages of 83% or above and Kappa-values of at least 0.66, showing ‘substantial’ agreement. 9 of 11 image schemas are matched to the test color pairs in at least 90% of the trials, resulting in ϰ-values of over 0.80 - an ’almost perfect’ agreement with the predicted image-schematic color mappings. Two one-sample *t*-tests against chance level ($\varkappa = 0.00$) yield a significant difference from the test condition (M = 0.84, SD = 0.22), $t(10) = 12.61$, $p < .001$, $d = 1.16$, but not for the control condition (M = -0.04, SD = 0.10), $t(10) = -0.36$, $p = .724$. In addition, a paired-samples t-test yielded a significant difference between the averaged $\varkappa$-values of test and control condition across all investigated image schemas, $t(10) = 8.77$, $p < .001$, $d = 5.15$, as well as between the averaged percentages of test (M = 91.91, SD = 3.38) and control condition (M = 48.18, SD = 4.91), $t(10) = 8.79$, $p < .001$, $d = 10.37$.

Next, the predicted absolute distances between the sample colors of a color pair based on the results of Study 1 (German data) and Study 2 (Japanese data) and the $\varkappa$-values from the Japanese subjects of Study 4 were correlated. The Pearson correlation coefficients indicated strong positive relationships for Japanese, $r(17) = .758$, $p < .001$, $r^2 = .575$, as well as German data, $r(17) = .672$, $p = .003$, $r^2 = .452$. This is not surprising as both predicted distances also correlate strongly with each other, $r(81) = .818$, $p < .001$.

In order to compare the results of the German (Study 3) and Japanese subjects (Study 4) regarding their performance in assigning color pairs of test and control condition to image-schematic dimensions, a Pearson correlation coefficient was calculated between the corresponding $\varkappa$-values, yielding a strong positive relationship, $r(11) = .917$, $p < .001$.

5.4.3. Discussion

Similar to Study 3, BIG-SMALL assignments were much less reliable that those of any other image schema. As already discussed in sec. 5.3.3, this might be due to the background/foreground contrast ratio of the stimulus material as well as due to the presentation of two same size color patches next to each other.

The test color pairs of 9 out of 11 image schemas were assigned in at least 90% of the trials to the image schema dimension in a way predicted by the position of the individual colors on the image schema dimension based on the results of Study 1 and Study 2, and can be recommended as color population stereotypes. Basing the predictions for Japanese subjects on German data resulted in a loss of 12.3% of the explained variance. Thus, even when the prediction is based on data obtained from a different cultural sample, the correlation between the prediction and empirical data was nevertheless 'strong', supporting the universality claim of image-schematic color associations. Moreover, the the assignments of Germans and Japanese in Study 3 and Study 4 strongly correlate, providing additional evidence for cross-culturally stable color population stereotypes. Similar to what has been observed in Study 3, test color pairs were non-randomly assigned to image schemas, whereas the control color pairs were assigned at chance level. These results show that, on average, the greater the distance between two (image schema-relevant) colors on an image schema dimension, the more reliable participants can identify the underlying mapping. By gathering participants' image schema ratings on single color patches, it was possible to predict their dimensional assignment of an image schema to a pair of colors. Moreover, this prediction was also suitable to hypothesize about the color assignment of another culture, since the data of both German and Japanese subjects does not considerably vary.

5.5. Chapter 5 Recap

In this chapter, 16 associations between image schemas and colors received empirical investigation. In four online surveys, the relationship between hue, saturation and brightness and the image schemas BIG-SMALL, BRIGHT-DARK, CLEAN-DIRTY, FAR-NEAR, FAST-SLOW, FULL-EMPTY, HARD-SOFT, HEAVY-LIGHT, LOUD-SILENT, OLD-YOUNG, PAINFUL-NOT PAINFUL, SMELLS GOOD-SMELLS BAD, SMOOTH-ROUGH, STRONG-WEAK, TASTES GOOD-TASTES BAD and WARM-COLD were examined. In Study 1, German participants matched single color patches to 16 image schema dimensions. For each image schema, a regression model was derived that quantifies the impact of hue, saturation and brightness on the healthy observers' judgement. These empirical results were discussed in light of four theories or frameworks that can be applied to explain and predict effects in the field of color psychology.

First of all, it was noted that image schema-color associations do not fit into the scope of color-in-context theory (Elliot & Maier, 2012a). Although the framework makes useful high-level assumptions about when color exhibits its influence, it is difficult to break these statements down to concrete predictions. On the contrary, some viable predictions can be deducted from Affective Meanings Systems (Osgood, 1969). Affective Meanings Systems predicts affective associations between concepts with a high dimensional overlap in the three-dimensional space of meaning. Since Osgood and colleagues already gathered data on many concepts including some hues and image schemas from a variety of cultures, predictions regarding image-schematic color relationships could be derived, although somewhat laboriously. In total, 23% of the empirical results of Study 1 were correctly predicted following this approach. A database extension for the color attributes saturation and brightness, as well as several image schemas, would probably increase the predictive power of the theory for color psychology. Major shortcomings of Affective Meanings Systems as a theory of color psychology are the separate investigation of hue, saturation and brightness, the underspecified way of how to compare and weight different EPA-scores, and the fact that the current database on color values is based on only linguistic stimuli, rendering the overall explanatory and predictive power unsatisfactory. As a third theoretical approach that might be fruitful for color psychology, the Ecological Valence Theory on Human Color Preferences (Palmer & Schloss, 2009), was considered. According to Schloss et al., people prefer objects in their favorite color, and probably recall these associations when making judgements as in Study 1. As the individual learning history varies between individuals (although there may be general trends), no precise predictions could be derived without having to conduct further studies. Moreover, according to the subjects' feedback, color preferences played a minor role. Finally, CMToC was applied to the research outcomes of the first study. Following this framework, the physical, biological or statistical origin of the association have to be (at least hypothetically) determined. After that, predictions can be deducted regarding the influence of color characteristics on different image schema dimensions in a given context, but also regarding the direction, strength and cultural stability. Overall, almost 80% of the uncovered image schema-color associations were correctly predicted. Some of the 'new' image schemas (DIRTY-CLEAN, LOUD-SILENT, SMELLS GOOD-SMELLS BAD) worked equally well as more established ones, while others (PAINFUL-NOT PAINFUL, TASTES GOOD-TASTES BAD) did not correlate with color attributes. It is concluded that the explorative character of the new image schemas should stimulate further research to gather empirical data on their existence as they are promising candidates that extend the sensorial range of image schemas. Moreover, CMToC inspires hypothesizing of why some associations are not in line with the predictions made by literature, sparking further research. Therefore, it was concluded that CMToC is the most promising theoretical approach for predicting semantic color associations with physical properties, as studied in the first four questionnaires.

Another focus was to examine whether similar results can be obtained in a different culture and how well the data of Japanese subjects can be predicted from regression models derived from data of German participants. Therefore, the first study was replicated with Japanese subjects (Study 2). In this independent test of the theory, 73% of the empirical image schema-HSB relationships were correctly predicted by CMToC. Using the regression models of the German participants, a significant proportion of the observed outcomes of the same survey conducted with Japanese subjects could be explained for all image schemas. Generally, the image-schematic ratings between both subject groups were highly similar, with the three image schemas HARD-SOFT, LOUD-SILENT and PAINFUL-NOT PAINFUL having a considerably weaker overlap than the other 13. The predictions of color-in-context theory, Affective Meanings Systems and Ecological Valence Theory on Human Color Preferences regarding cultural stability and variability of image schematic-color associations are overall weak. Affective Meanings Systems states that there are universal trends in in affective meanings of color (Adams & Osgood, 1973). Following this, if Japanese and Germans evaluate image schemas with similar EPA-scores, the overlap between image schemas and colors should be similar. Although this approach can make specific predictions when similarities can be expected, it does not explain where a potential difference in evaluating concepts stems from. CMToC, on the contrary, makes predictions for each image schematic-color association based on the level of prior knowledge on which the association is potentially located. Innate associations and those acquired on a sensorimotor level of prior knowledge are shared across cultures. Color associations are expected to vary only if they are based on correlations in the man-made environment. This classification allows to pinpoint color associations that are subject to intercultural and intracultural influences, and also to identify or reason about sources of origin. Therefore, it is concluded that CMToC provides a more compelling approach compared to Affective Meanings Systems although at the cost of being less parsimonious.

Study 3 and 4 aimed at confirming the findings of the previous studies by extrapolating single color judgements to two-color combinations in an effort to unveil cross-cultural color population stereotypes. German and Japanese subjects matched color pairs that differed in their distance on the image schema dimension to image-schematic end poles. Both experiments show that the greater the distance of individual colors on the image schema dimension, the more reliable and non-random the assignment. Again, subjects of both cultural backgrounds made very similar assignments. By using the data derived from German participants compared to data derived from Japanese subjects, a loss of explained variance in the prediction of 12.3% was observed. Therefore, when high accuracy is need, it is recommended to generate color regression models for a specific culture. However, the models developed in this work are sufficient to predict a substantial part of semantic color associations with physical properties across cultures and can be used to save effort to generate new more culture-specific models.

The empirical findings of the first four studies not only advance our theoretical understanding of semantic associations between color and other physical properties, but also help designers in making appropriate color choices. When physical properties like those represented by the 16 image schemas investigated in this work play a role in user interfaces, designers can now select colors that best match with these properties. For example, WARM-COLD should be aligned with red-blue hues regardless of brightness and saturation, whereas a pleasant taste goes along with high brightness, and for HEAVY-LIGHT, all three color attributes have to be considered. Moreover, based on the regression models, it is possible to compare two or more colors regarding their associative strength with an image-schematic dimension. By relying on color associations with image schemas and thus prior knowledge on the sensorimotor level, designs are likely understood by a larger audience compared to arbitrary relations or associations based on cultural and expertise levels of prior knowledge. Although the online surveys were conducted on a variety of different display devices with probably widely varying color presentation profiles, the internal validity as determined by split-half reliability between different raters was very high. Moreover, this method enhanced the ecological validity of the studies as it is closer to real-world settings than highly controlled laboratory conditions. Still, additional empirical work in more applied contexts is required. By following the approach of CMToC, designers are able to identify some boundary conditions of color associations in specific settings, e.g. with different background contrasts. As not only physical concepts but also abstract content play a major role in user interfaces, color associations with more abstract concepts need to be elaborated. After this to date most comprehensive cross-cultural survey of image schemas and related color associations, the next chapter aims at investigating if the image schematic-color associations do expand to metaphorical extensions of those image schemas, forming color substituted image-schematic metaphors.

6. Conveying Abstract Concepts with Color

When abstract information is represented in a user interface, designers have to make appropriate color choices, as unsuitable colors are likely to hamper intuitive use. In this chapter, population stereotypes (Bergum & Bergum, 1981) in form of color substituted image-schematic metaphors are derived from CMToC and put to empirical test. If a large proportion of German and Japanese subjects matches colors to abstract domains in a similar way, practical rules for the design of user interfaces regarding abstract information can be derived that facilitate intuitive use.

How does CMToC predict color associations with abstract concepts? Like demonstrated in the previous chapter, cross-domain links between colors and image schemas exist. And since many abstract concepts are perceptually grounded in image schemas, it is possible that the associations between image schemas and colors do expand to metaphorically related abstract concepts. Following this idea, people should match colors to metaphorical extensions of image schemas based on how they matched colors to image schemas (Löffler, 2014b). For example, as the WARM-COLD image schema is related to hue (e.g., RED IS WARM - BLUE IS COLD), metaphorical extensions of WARM-COLD should also be linked to hue similarly. One way to determine metaphorical extensions as demonstrated in sec. 4.4.1 is by literature study and language analysis. As expressed in terms like *I hold warm feelings towards her* or *He gave her the cold shoulder*, we do use "warm" and "cold" to talk about the less tangible concept of INTIMACY (INTIMATE IS WARM - DISTANT IS COLD). By combining the relationship of image schema and color with the image-schematic metaphor, the color substituted image-schematic metaphor INTIMATE IS RED - DISTANT IS BLUE results, linking both color and abstract content. When image schemas are instantiated by a color with which they naturally correlate, they should meet the subconscious expectations of the participants (P. Walker et al., 2010), which should in turn result in more intuitive interaction (Hurtienne & Blessing, 2007; Hurtienne, 2011).

In order to test this idea, candidate colors were derived from the regression models developed in the previous chapter. Then, an experiment was conducted that tests whether the perceived color of objects influences conceptual judgements on abstract concepts that are metaphorically linked to the underlying image schema (color substituted image-schematic metaphors). Focussing on metaphorical extensions of the image schemas BIG-SMALL, HEAVY-LIGHT, STRONG-WEAK and WARM-COLD, German and Japanese subjects matched colored objects to abstract words. From the results color population stereotypes are derived that are “beneficial in that they shorten the time for learning and practicing, that they enhance user performance, and that they reduce the number of erroneous actions, especially under stressful conditions” (Hurtienne et al. 2009, p.61). Contrasting to the typical approach of investigating population stereotypes ad-hoc, CMToC enables designers to derive new design rules for conveying abstract information with color that can be justified on a theoretical basis.

6.1. Study 5 - Color Substituted Image-Schematic Metaphors in German and Japanese Subjects

Studies on population stereotypes involving image schemas and colors were mostly conducted to inform the design of TUIs and facilitate intuitive interaction. For example, findings of Hurtienne et al. can be interpreted as promising first evidence that abstract concepts are systematically linked to specific color attributes via image-schematic metaphor. Even though not in focus of their research, the authors investigated if the manipulation of the brightness (instantiated as *black* and *white* or *dark* and *light* shades of the same color) of tangible objects is metaphorically linked to abstract concepts such as *happiness* or *religiousness.* Moreover, it was tested whether color temperature (instantiated as *red* vs. *blue*) is metaphorically linked to abstract concepts like *emotionality* or *difficulty* (Hurtienne et al., 2009). The results show that the majority of the metaphors works well with the color manipulations. While in some cases, the color condition even performed better than a haptic instantiation of the image schema (WARM-COLD), others performed worse (different shades of the same color for BRIGHT-DARK mappings). Having a closer look at the stimulus material, a possible explanation for this bad performance in the BRIGHT-DARK condition is that the color attribute saturation rather than brightness was manipulated (Löffler, 2014b). Overall, the findings suggest that color-for-haptic substitutions might be a viable approach worth further exploration.

Two more recent studies explicitly investigated whether substituting color for haptic attributes in physical-to-abstract mappings could be a feasible alternative for the design of tangibles. The argument of these studies is that changing colors in TUIs is a less expensive and more practical way to dynamically change, for example, the size, weight or temperature of tangibles (Löffler, Paier, Toriizuka, Ikeda, & Hurtienne, 2015; Löffler et al., 2016). Basically, subjects in both experiments had to indicate which one of a pair of colored objects matches a given metaphoric expression best. Following this procedure, color-for-haptic substitutions were tested for their effectiveness with 12 (Löffler et al. 2015) and 15 conceptual metaphors (Löffler et al. 2016) of the BIG-SMALL, HEAVY-LIGHT and WARM-COLD image schema. In both studies, two thirds of the substitutions were identified by the majority of the participants and subsequently recommended as design guidelines, e.g. DIFFICULT IS LOW SATURATION.

Löffler et al. 2016 employed a more sophisticated experimental design, allowing to compare the matching performance of the participants in four different conditions that differed in the way in which the image schema was operationalized: haptically, via color, combining haptic and visual (color) instantiation in a congruent and in an incongruent way. The authors conclude that "color can replace haptic attributes in metaphoric mappings and that designers need to explicitly design for color, because metaphor-incongruent colors can hamper the effectiveness of metaphorical mappings" (Löffler et al. 2016, p.118). In total, 81% of the color-for-haptic substitutions were identified by the participants.

Although these studies offer promising evidence that colors are not only associated with image schemas but also extend to metaphorically related abstract domains, they also face a number of shortcomings and leave open questions for future research that will be addressed in Study 6 presented in this chapter. Löffler et al. 2015 altered the object colors via Augmented Reality. Because of using this technology, color changes were limited to hue or saturation, because changes in brightness would have negatively affected the three-dimensional impression of the objects. However, especially the brightness dimension plays an important role in linking semantic content to color (Adams & Osgood, 1973; Valdez & Mehrabian, 1994; Crozier, 1997; Hupka et al., 1997). The work published by the research group in 2016 altered all color attributes. Besides minor drawbacks in the stimulus material of this study like the use of white control objects (*white* is often associated more with one image schema pole than the other), the study leaves some questions open. For example, whether similar results can be obtained in different cultures, if other measures of intuitive use are also affected by the choice of colors, and if the results extend to other image schemas as well. Moreover, an aspect worth investigating is the impact of the modality in which the image schema instantiation is perceived (e.g., visual, haptic, or multimodal). If the user is engaged in a primary task that involves one modality (e.g. haptics), the interaction possibilities for a second task can be limited to other modalities not involved, e.g. vision. It is therefore important to find out whether color substituted image-schematic metaphors can be identified with limited interaction capabilities, or if a multimodal perception is necessary for optimal results (i.e., inspecting objects by touch and vision).

In order to confirm the effects of color on abstract conceptual judgements for different cultures, a conceptual replication of the work of Löffler et al. 2016 is performed with German and Japanese subjects. The experimental procedure is complemented by measuring reaction times and eliciting preference data as measures of efficiency and satisfaction, adding metaphorical extensions of the haptic target domain STRONG-WEAK, and investigating whether image-schematic metaphors are similarly identified when participants look and touch tangible objects, or only use one of these modalities. As Löffler et al. 2016 demonstrated that it is not specific color variants that drive the effectiveness of the mappings, only one color variant per image schema will be used in Study 5 to keep the length of the experiment reasonable. It is expected that subjects of both cultures prefer metaphorically congruent objects to metaphorically incongruent objects, no matter if the source domain is instantiated haptically or visually through color and if the mapping is perceived uni- or multimodal. This preference should be reflected in all three measures of intuitive use: metaphor-consistent choices, reaction times and preference ratings. The identification is expected to be impaired if color and haptic property send opposing signals. Details of the experimental procedure and material followed by the derivation of concrete hypotheses are described in the next paragraphs.

6.1.1. Method

6.1.1.1. Participants

28 German students from the University of Würzburg, Germany, and 47 Japanese volunteers from Nihon University, Japan, were recruited and participated for course credit or 11 Euro. The participants' ages ranged from 19 to 55 years (M = 22.3, SD = 5.33). All German participants were German native speakers, all Japanese subjects had Japanese as their mother tongue. All participants had no reported defective color vision and passed four plates of the Ishihara Color Test (Ishihara, 1917) with maximum one error. Their favorite colors were green (20.0%), blue (18.7%), red (14.7%), orange (12.0%), yellow (9.3%), black (5.3%), white (5.3%), violet (2.7%), and other (8.0%). The subjects were all naïve as to the subject under investigation.

6.1.1.2. Procedure

Before starting with the main experiment, participants completed a demographic questionnaire. After that, they were instructed that in each experimental trial they will be presented with two objects and an adjective describing an abstract concept (e.g., *powerful*). Their task was to inspect the objects and select the object that best matched the given adjective. Depending on condition, participants inspected the objects either through looking at them while lifting them with both hands (conditions 1-4), or through lifting them with both hands while being blindfolded (condition 5), or by only visual inspection (condition 6). The details of each condition are described in the section about the stimulus material. The adjective was printed in black letters on white paper and placed together with two objects on a wooden tray in front of the participants. In the condition in which the participants were blindfolded, the adjective was read out loud by the experimenter.

No time limit was imposed on the participants, but the time participants needed to give their answers was recorded. The recording started when subjects had both objects in their hands and stopped when the chosen object was placed on the table. Note that in the experimental condition in which the participants were not allowed to touch the objects (condition 6), time recording started when the objects were placed on the table and stopped when the participants pointed to one object. To avoid sequence effects, the objects and adjectives were presented in random order. After finishing all trials, the participants were asked to fill out a questionnaire of how suitable certain object sets were in representing certain metaphors. Following this, they had to provide qualitative feedback on their strategies and thoughts during the experimental procedure. The whole experiment lasted about 45 minutes, with no considerable variance between the participants. The Japanese subjects participated in the experiment in July 2014, whereas the German subjects took part in the experiment using the same stimulus material in August 2014. The laboratories in which the experiment took place had covered windows and were equipped with daylight lamps (about 220 Lux). The room temperature for both rooms was about 21 degrees with 40% humidity.

6.1.1.3. Material

Stimuli

Image Schema Instantiations and Experimental Conditions Four pairs of objects were created for each of the image schemas BIG-SMALL, HEAVY-LIGHT, STRONG-WEAK and WARM-COLD, making a total of 16 object pairs (see Tab. 6.1, first four rows). Pair A instantiated the image schema in a haptic way while keeping color constant at neutral gray. Pair B instantiated the image schema in a metaphoric color variant without haptic differences. Pair C instantiated the image schema both haptically and via color cues in a metaphorically congruent way, while pair D combined the color cues in a metaphorically incongruent way compared to the haptic cues. The chromatic colors were chosen to resemble image schemas and were based on the regression models of Study 1. While color-temperature associations were mainly influenced by hue, the main drivers for size-, strength- and weight-associations are saturation and brightness. Therefore, the hues in the size-, strength- and weight-conditions were kept constant. The object color was achieved by covering the objects in colored cardboard paper.

In detail, the objects had the following properties (HSB values in parentheses):

- BIG-SMALL:
  - A. differently sized cardboard boxes in neutral gray (n.d.,0,50) : 2.5x5.1x 5.1cm (BIG), 1.25x2.5x2.5cm (SMALL)
  - B. same sized cardboard boxes (1.87x3.75x3.75cm): dark red (0,100,25) (BIG), pink (0,25,100) (SMALL)
  - C. differently sized cardboard boxes: 2.5x5.1x5.1cm in dark red (0,100,25) (BIG), 1.25x2.5x2.5cm in pink (0,25,100) (SMALL)
  - D. differently sized cardboard boxes: 2.5x5.1x5.1cm in pink (0,25,100) (BIG), 1.25x2.5x2.5cm in dark red (0,100,25) (SMALL)

- HEAVY-LIGHT:
  - A. same sized cardboard boxes with different weights in neutral gray (n.d.,0,50): 200g (HEAVY), 14g (LIGHT)
  - B. same sized cardboard boxes with same weight (100g): dark violet (282,100,25) (HEAVY), violet (282,75,100) (LIGHT)
  - C. same sized cardboard boxes with different weights : 200g in dark violet (282,100,25) (HEAVY), 14g in violet (282,75,100) (LIGHT)
  - D. same sized cardboard boxes with different weights : 200g in violet (282,75,100) (HEAVY), 14g in dark violet (282,100,25) (LIGHT)
- STRONG-WEAK:
  - A. same sized magnets in neutral gray (n.d.,0,50): 4000g/39.2N (STRONG), 408g/4N (WEAK)
  - B. same sized magnets with same strength 2200g/21.6N: brown (37,100,50) (STRONG), bright brown (37,25,75) (WEAK)
  - C. same sized magnets: 4000g/39.2N in brown (37,100,50) (STRONG), 408g/4N in bright brown (37,25,75) (WEAK)
  - D. same sized magnets: 4000g/39.2N in bright brown (37,25,75) (STRONG), 408g/4N in brown (37,100,50) (WEAK)
- WARM-COLD:
  - A. two identical glass jars in neutral gray (n.d.,0,50) filled with water: 50.0°C (WARM), 5.0°C (COLD)
  - B. two identical glass jars filled with water at room temperature: orange (50,100,100) (WARM), blue (250,75,100) (COLD)
  - C. two identical glass jars filled with water: 50.0°C in orange (50,100,100) (WARM), 5.0°C in blue (250,75,100) (COLD)
  - D. two identical glass jars filled with water: 50.0°C in blue (250,75,100) (WARM), 5.0°C in orange (50,100,100) (COLD)

A manipulation check was conducted to test whether the haptic instantiations were suitable representations of the four image schemas. Ten subjects rated each object (pair A) on a 7-point Likert scale in terms of their main haptic attribute (BIG-SMALL, HEAVY-LIGHT, STRONG-WEAK or WARM-COLD). Four t-tests for dependent means indicated that the *warm* ($MW = 5.40$, $SD = 0.49$) and *cold* ($MW = 2.60$, $SD = 0.49$) glass jars were differently perceived regarding their temperature, $t(9) = 3.721$, $p < .005$, the *strong* ($MW = 5.70$, $SD = 0.46$) and *weak* ($MW = 2.80$, $SD = 0.98$) magnets were judged differently regarding their strength, $t(9) = 9.222$, $p < .001$. The *big* box ($MW = 4.10$, $SD = 0.94$) was rated as larger than the *small* ($MW = 2.10$, $SD = 0.70$) one, $t(9) = 6.708$, $p < .001$, and the weight of the *heavy* box (MW

= 4.80, SD = 1.08) was rated higher than that of the *light* box (MW = 1.30, SD = 0.46), $t(9)$ = 10.247, $p < .001$.The chosen colors were not tested in the manipulation check but are based on the results of Study 1.

Table 6.1.: *Overview of the Six Experimental Conditions, Four Object Pairs and Interaction Modalities for each Image Schema*

			Image schema			
Condition	Object Pair	Modality	BIG-SMALL	HEAVY-LIGHT	STRONG-WEAK	WARM-COLD
(1) Haptic	A	look and touch	big small	heavy light	strong weak	warm cold
(2) Color	B	look and touch	same size	same weight	same strength	same temperature
(3) Congruent	C	look and touch	big small	heavy light	strong weak	warm cold
(4) Incongruent	D	look and touch	small big	light heavy	weak strong	cold warm
(5) Haptic (blindfolded)	A	touch	big small	heavy light	strong weak	warm cold
(6) Color (only looking)	B	look	same size	same weight	same strength	same temperature

Six experimental conditions were created for each image schema, all of which are depicted in Tab. 6.1. In the first condition, the object pairs described under A., only differing in a haptic property but not in color, were inspected by the participants both visually and haptically (this condition is referred to as *haptic* condition). The different sized boxes of the BIG-SMALL image schema are an exception, since size is perceived visually as well as haptically. For the second condition (referred to as *color* condition) the objects described under B., differing in color but not in the haptic property, were visually and haptically inspected. In the third condition, called *haptic-color congruence* condition, participants visually and haptically inspected the metaphorically congruent object pairs described under C. and in the fourth condition, consisting of the metaphorically incongruent pair D., visual and haptic inspection was encouraged as well (*haptic-color incongruence* condition). In condition five (*blindfolded*), the object pairs described under A. were used, and the participants were blindfolded, being only able to inspect the items haptically. Finally, in condition six (*only looking*), the objects described under B. were used, but participants were only allowed to have a look at the objects without touching or lifting them.

Metaphoric Expressions The following metaphoric expressions for the image schemas BIG-SMALL, HEAVY-LIGHT, STRONG-WEAK and WARM-COLD were determined by literature research as well as own analyses. In addition to well documented metaphorical extensions, sentences in three languages (German, Japanese, English) were collected containing the image schema itself (e.g. *big, small*) and synonyms like *huge* or *tiny* in their metaphorical, not literal, meaning. Sentences were collected from the online dictionaries *Digitales Wörterbuch der deutschen Sprache* (digital dictionary of the German language) [1], *Jisho* (Japanese-English dictionary) [2], and The Free Dictionary [3]. The sentences were clustered according to their meaning. Two experts on image schemas and CMT selected the five most stable metaphors. Selected metaphors along with example sentences in German, Japanese and English, if applicable, are depicted in Tab. 6.2.

[1]http://www.dwds.de, last accessed on 13.08.2016
[2]http://jisho.org, last accessed on 13.08.2016
[3]http://www.thefreedictionary.com, last accessed on 13.08.2016

Table 6.2.: *Metaphorical Extensions of the* BIG-SMALL *Image Schema*

Metaphor	German	Japanese	English
KNOWING IS BIG - UNKNOWING IS SMALL (TOLAAS, 1991)	Den eigenen Horizont erweitern. (Widening the own horizon.) Du bist total beschränkt. (You are very narrow-minded.)	あなたは興味の対象を広げる必要がある. (You have to broaden your horizons.) 彼の見識は狭い. (His insights are narrow.)	Traveling will widen your horizon. They are narrow-minded people.
SIGNIFICANT IS BIG - INSIGNIFICANT IS SMALL (LAKOFF & JOHNSON, 1980)	Eine große Persönlichkeit in der Geschichte. (A big personality in history.) Eine kleine Lüge. (A small lie.)	その事件は，我が社にとって重大な影響をもたらす. (The incident has a big impact on our company.) 私は小さな変化を見逃さなかった (I did not overlook even small changes.)	He's a giant among writers. It was only a small crime.
IMPORTANT IS BIG - UNIMPORTANT IS SMALL (LAKOFF, 1994)	Eine große Entdeckung. (A big discovery.) Er hält sich an kleinen Details auf. (He is held back by little details.)	彼はじゅうだい重大なはっけん発見をした. (He made a big discovery.) 小さい過ち. (A small error.)	Let's look at the larger issues. He concentrates on little details.
POWERFUL IS BIG - POWERLESS IS SMALL (Baldauf, 1997)	Die großen Persönlichkeiten unserer Zeit. (The big personalities of our time.) Das Anliegen der kleinen Leute. (The concern of the small [ordinary] people.)	彼は偉大な人物だ. (He is a big [great] person.) 彼は小さく見える. (He looks small [powerless].)	He's a big man in industry. The federal union – a political dwarf.
MORE IS BIG - LESS IS SMALL (TOLAAS, 1991)	Die Klasse ist größer als die andere. (This class is bigger than that one.) Mein Anteil ist viel kleiner als deiner. (My share is much smaller than yours.)	このクラスはあのクラスより大きい. (This class is bigger than that one). 8より小さい数. (A number smaller than 8).	This class is bigger than that one. I have a smaller amount of money than he does.

Table 6.2.: *Metaphorical Extensions of the* HEAVY-LIGHT *Image Schema*

Metaphor	German	Japanese	English
IMPORTANT IS HEAVY - UNIMPORTANT IS LIGHT (Baldauf, 1997)	Ein schwerwiegendes Urteil. (A weightily decision.) Eine leichte Schnittwunde. (A light cut.)	当局はその事件を重く見た. (The authorities took the incident very heavy.) 軽い病気. (A light illness.)	Heavy matters of state. Light, idle chatter.
SAD IS HEAVY - HAPPY IS LIGHT (Stefanowitsch, 2006)	Eine schwere Depression. (A heavy depression.) Er war erleichtert. (He was unburdened.)	今日はどうも頭が重い. (Today my head feels heavy [depressed].) 心も軽くピクニックに出掛けた. (I leave for the picnic with a light heart [cheerfully]).	Heavy sadness. With a light heart.
MORE IS HEAVY - LESS IS LIGHT (LAKOFF, 1994)	Ein schwerer Trinker. (A heavy drinker.) Er spricht mit leichtem Akzent. (He talks with light accent.)	重い税. (Heavy taxes.) 軽い罰. (Light punishment.)	He's a heavy smoker. She was always a light drinker.
DIFFICULT IS HEAVY - EASY IS LIGHT (GIBBS ET AL., 2004)	Eine schwere Geburt. (A heavy birth.) Eine leichte Aufgabe. (A light task.)	Only EASY IS LIGHT: 横綱は相手を軽く打ち負かした. (The Yokozuna (sumo champion) bet his opponent lightly.)	She's weighed down by responsibilities. Light duties.
GUILTY IS HEAVY - INNOCENT IS LIGHT (TOLAAS, 1991)	Die Schuld lastet auf ihm. (The burden of guilt rests on him.) Sie konnte ihr Gewissen erleichtern. (She could unburden her conscience.)	罪が重い. (A heavy crime). 私の潔白が証明されて，心が軽くなった. (My mind became light since my innocence was proved.)	The burden of the blame. A weight fell from her.

Table 6.2.: *Metaphorical Extensions of the* STRONG-WEAK *Image Schema*

Metaphor	German	Japanese	English
MORE IS STRONG - LESS IS WEAK (Baldauf, 1997)	Ich habe starkes Interesse daran. (I have strong interest in that.) Das ist nur ein schwacher Trost. (It brings only weak [little relief.)	彼は責任感が強い. (He has a strong [keen] sense of responsibility.) 彼の主張は根拠が弱い. (He has weak basis for his claim.)	An army 20,000 strong. A weak market for oil stocks.
POWERFUL IS STRONG - POWERLESS IS WEAK (Baldauf, 1997)	Die großen Verlage sind damit noch stärker geworden. (Given that, the big publishers got even more strong.) Eine schwacher Herrscher. (A weak ruler.)	気の強い娘. (A strong-minded girl.) 説得力の弱い議論. (A weak [unconvincing] argument.)	A strong leader. A weak argumentation.
INTENSE IS STRONG - LESS INTENSE IS WEAK	Was riecht hier so stark? (What is this strong odor?) Eine schwache Beleuchtung. (Weak lighting.)	強い光. (A strong [intense] light.) 弱い風. (A weak breeze.)	A strong odor of burning rubber. A weak tea.
COMPETENT IS STRONG - INCOMPETENT IS WEAK (Baldauf, 1997)	Das war total stark! (That was really strong [great].) Eine schwache Leistung. (A weak performance.)	歴史に強い. (He is strong in history.) 彼は数学に弱い. (Mathematics is his weak point.)	Their chief strength is technology. He is weak at football.
MALE IS STRONG - FEMALE IS WEAK	Männer - das starke Geschlecht. (Men - the strong sex.) Das schwache Geschlecht. (The weak sex.)	No metaphor in Japanese.	Are men the strong sex? This revolution lifted the weak sex onto an equal level as that of men.

Table 6.2.: *Metaphorical Extensions of the* WARM-COLD *Image Schema*

Metaphor	German	Japanese	English
INTIMATE IS WARM - DISTANT IS COLD (Lakoff & Johnson, 1999)	Wir sind schnell warm geworden. (We got warm with each other quickly.) Sie empfing ihn kalt. (She welcomed him coldly.)	二人は熱い仲だ. (The relationship between those two is hot.) 何で二人の仲は冷たくなったのか. (What cooled the relationship between the two?)	We got warm with each other. The relationship cooled off.
PROBLEMATIC IS WARM - UNPROBLEMATIC IS COLD (Baldauf, 1997)	Eine hitzige Debatte. (A hot debate.) Wir haben die Sache auf Eis gelegt. (We put the topic on ice.)	No metaphor in Japanese.	UN presence in the territories would only serve to inflame the situation. The project was put on ice.
EMOTIONAL IS WARM - UNEMOTIONAL IS COLD (TOLAAS, 1991)	Die Leidenschaft zwischen uns brennt immer noch. (The passion between us is still burning.) Ein eiskalter Mörder. (An ice-cold murderer.)	心の暖かい人. (A warm-hearted person.) 心の冷たい人. (A cold-hearted person.)	A warm personality. Cold logic.
ACTIVE IS WARM - INACTIVE IS COLD (SAMBRE, 2000)	Der Konfikt ist wieder entbrannt. (The conflict flared up.) Die Löhne einfrieren. (Freezing the wages.)	彼は将来の計画を熱い調子で語った. (He told me about his future plans with hot (熱い.) mood [in an excited tone]. 利率が凍結された. (Interest rates were frozen (凍.).)	The movement for democracy began to heat up. Frozen rents.
HAPPY IS WARM - SAD IS COLD (Stefanowitsch, 2006)	Freude, schöner Götterfunken. (Joy, fair spark of the gods.) Oh wie ist es kalt geworden [Song of von Fallersleben]. (Oh, how cold has it gotten.)	幸せは人の温かさです. (Happiness is human warmness.) 悲しくて身も心も凍えそうだ. (I am sad as if both my body and soul would be frozen.)	Warm joy. A deep cold sadness overtook me.

The following adjectives were chosen to represent metaphoric expressions of the image schemas BIG-SMALL, HEAVY-LIGHT, STRONG-WEAK and WARM-COLD and shown to the participants:

- BIG-SMALL: important-unimportant, significant-insignificant, powerful-powerless, more-less, knowing-unknowing

- HEAVY-LIGHT: important-unimportant, sad-happy, more-less, difficult-easy, guilty-innocent
- STRONG-WEAK: more-less, powerful-powerless, intense-less intense, competent-incompetent, male-female
- WARM-COLD: emotional-unemotional, active-inactive, intimate-distant, problematic-unproblematic, happy-sad

Note that this stimulus material is identical to the one employed by Löffler et al. 2016 with the exception of the newly added image schema STRONG-WEAK and changes in two temperature-metaphors (INTIMATE IS WARM - DISTANT IS COLD instead of FRIENDLY IS WARM - UNFRIENDLY IS COLD, PROBLEMATIC IS WARM - UNPROBLEMATIC IS COLD instead of FAMILIAR IS WARM - UNFAMILIAR IS COLD).

Two lists of adjectives were prepared. On list 1, one word from each of the 20 adjective pairs was placed (e.g. *knowing*), and its opposite was placed on list 2 (e.g. *unknowing*). The lists were balanced with regard to the number of positive and negative words they contained. Half of the participants were presented with list 1, the other half with list 2. Therefore, each participant matched 20 adjectives (metaphors across all four image schemas) to all object pairs in the six conditions, resulting in 120 trials per session.

Preference Questionnaire The final questionnaire which participants were required to fill out after they completed all matching trials is shown in sec. A.5.1. Five versions were created (containing one metaphor for each image schema) and each version was handed out to one fifth of the participants. The questionnaire showed all four object sets as instantiations of an image schema (16 sets in total). For each object set related to an image schema, the subjects were given the two opposing adjectives resembling a metaphorically related abstract concept. Their task was to rate each object set regarding how well it represented the metaphor on a 5-point Likert scale from (– = very ill suited to ++ very well suited).

Auxiliary data To gather demographic data of the participants, an online survey was created using EFS Survey and translated into German as well as Japanese. In this online survey, participants completed four plates of the Ishihara Color Test (Ishihara, 1917). The plate numbers 1, 7, 8 and 17 were used to test for red-green visual impairments. This was followed by a questionnaire to be completed with age, gender, current occupation, highest education, ethnicity and mother tongue.

6.1.2. Hypotheses

The independent variables in Study 5 are

- the congruency of the colors to the haptic property (*haptic-color congruence* vs. *haptic-color incongruence* condition),
- whether participants could inspect the objects visually and haptically or only visually or haptically (*haptic* and *color* vs. *blindfolded* and *only looking* condition) (interaction modality),
- the instantiation type of the image schema that communicated metaphor-relevant information: in the *haptic* condition, only the haptic properties of weight, size, temperature and strength communicated metaphor-relevant information; in the *color* condition, the haptic properties were kept constant, so only the different colors conveyed potentially metaphor-relevant information; in the *haptic-color congruence* condition, participants can rely on both *haptic* and *color* information when making judgements in the metaphorical dimension (*haptic* vs. *color* vs. *haptic-color congruence* condition), and
- the participants' culture (German vs. Japanese).

Intuitive interaction as dependent variable was operationalized as

- Effectiveness: the percentage of object choices that were consistent with the metaphor (either as indicated by the haptic or color property, or in case of the *haptic-color incongruence* condition, as indicated by the haptic property),
- Efficiency: the time participants need to make their judgements, as well as
- Satisfaction: the participants' preference ratings regarding how suitable the different objects pairs are in resembling certain metaphors.

The hypotheses of this study as summarized in Tab. 6.3 are:

- (a) Effectiveness and efficiency of the mapping are impaired if color and haptic property send opposing signals compared to aligned signals, because image schema-color associations extend to metaphorically related abstract concepts. Participants prefer metaphorically congruent objects to metaphorically incongruent objects, no matter if the source domain is instantiated haptically, visually through color, or both.
- (b) The effectiveness of unimodal mappings is higher when they are accessed with two rather than one modality (Bertelson & Gelder, 2004). The efficiency of matching the objects to the adjectives is speeded-up when vision is involved, and slowed down when touch is involved.
- (c) Haptic instantiations of image schemas are more reliably matched to related metaphors than pure color instantiations, because the haptic experiences of weight, size, temperature and strength are the core of the sensory origin of the image schemas investigated in Study 5. Choices in the *color* condition

still differ from chance, as colors that indicate image schemas bias judgements in domains linked by metaphor. Moreover, redundant multimodal coding as in the *haptic-color congruence* condition conveys information more effectively than receiving information in only one modality (Bertelson & Gelder, 2004).

- (d) The results regarding the effectiveness and efficiency of and satisfaction with the mapping do not differ between German and Japanese subjects. Even if identical linguistic expressions for a specific metaphor are absent in one language, it is likely that the underlying conceptual image-schematic metaphor can be easily identified, as linguistic expressions of the same conceptual metaphor sometimes vary between cultures (Kövecses, 2005).

Table 6.3.: *Summary of the Hypotheses of Study 5*

	Matching Consistency	Time	Preference
(a)	congruent > incongruent	congruent < incongruent	haptic > incongruent
			color > incongruent
			congruent > incongruent
(b)	haptic > blindfolded	haptic < blindfolded	
	color > only looking	only looking < color	
(c)	congruent > color		
	haptic > color		
	congruent > haptic		
	color ≠ chance		
(d)	Japanese = German	Japanese = German	Japanese = German
	Japanese = German for Metaphor $_{\text{only in German}}$		

6.1.3. Data Analysis and Results

The combined results for both German and Japanese subjects are shown in Tab. 6.4 (separate values in Tab. A.9 and Tab. A.10, respectively). The different columns show the participants' percentage of answers consistent with the image-schematic metaphor and Cohen's kappa values as an index of metaphor strength (ϰ). Note that due to an error in the experimental material, matching consistency data and reaction times for *important-unimportant* in the BIG-SMALL image schema were not collected. Mean times in seconds for the matching process across all six conditions and per image schema are shown in Tab. 6.5 together with the preference ratings (-2 to +2) across the four instantiations per image schema.

Table 6.4.: *Metaphor Consistency in the Answers Across Four Presentation Styles in Six Conditions (German and Japanese Subjects)*

	Haptic		Color		Congruence		Incongruence		Blindfolded		Looking	
	%	ϰ	%	ϰ	%	ϰ	%	ϰ	%	ϰ	%	ϰ
Metaphor (... is big - ... is small)	**85**	**0.70**	**86**	**0.72**	**89**	**0.78**	**65**	**0.30**	**85**	**0.70**	**80**	**0.60**
significant - insignificant	86	0.72	93	0.86	89	0.78	61	0.22	88	0.76	83	0.66
important - unimportant	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
powerful - powerless	85	0.70	90	0.80	94	0.88	58	0.16	89	0.78	92	0.84
more - less	99	0.98	86	0.72	99	0.98	86	0.72	87	0.74	79	0.58
knowing - unknowing	74	0.48	75	0.50	72	0.44	54	0.08	74	0.48	61	0.22
Metaphor (... is heavy - ... is light)	**94**	**0.88**	**84**	**0.68**	**96**	**0.92**	**70**	**0.40**	**91**	**0.82**	**84**	**0.68**
important - unimportant	96	0.92	78	0.56	97	0.94	74	0.48	97	0.94	82	0.64
sad - happy	82	0.64	91	0.82	89	0.78	38	-0.24	75	0.50	93	0.86
more - less	100	1.00	81	0.62	99	0.98	96	0.92	96	0.92	81	0.62
difficult - easy	97	0.94	88	0.76	99	0.98	76	0.52	97	0.94	88	0.76
guilty - innocent	97	0.94	91	0.82	97	0.94	67	0.34	94	0.88	86	0.72
Metaphor (... is strong - ... is weak)	**91**	**0.82**	**86**	**0.72**	**92**	**0.84**	**74**	**0.48**	**84**	**0.68**	**88**	**0.76**
more - less	94	0.88	90	0.80	96	0.92	76	0.52	92	0.84	94	0.88
powerful - powerless	92	0.84	81	0.62	93	0.86	85	0.70	87	0.74	94	0.88
intense - less intense	87	0.74	79	0.58	91	0.82	82	0.64	78	0.56	83	0.66
competent - incompetent	90	0.80	78	0.56	85	0.70	76	0.52	78	0.56	69	0.38
male - female	90	0.80	94	0.88	96	0.92	58	0.16	88	0.76	96	0.92
Metaphor (... is warm - ... is cold)	**88**	**0.76**	**88**	**0.76**	**89**	**0.78**	**67**	**0.34**	**86**	**0.72**	**83**	**0.66**
intimate - distant	94	0.88	93	0.86	93	0.86	63	0.26	93	0.86	92	0.84
problematic - unproblematic	58	0.16	47	-0.06	61	0.22	61	0.22	64	0.28	36	-0.28
emotional - unemotional	97	0.94	97	0.94	99	0.98	61	0.22	99	0.98	93	0.86
active - inactive	97	0.94	96	0.92	97	0.94	61	0.22	86	0.72	93	0.86
happy - sad	93	0.86	96	0.92	98	0.96	67	0.34	87	0.74	100	1.00
Overall	**90**	**0.80**	**86**	**0.72**	**91**	**0.82**	**63**	**0.26**	**87**	**0.74**	**83**	**0.66**

Table 6.5.: *Mean Reaction Times (t) in Seconds and Preference Ratings (pref) form -2 to +2 Across Four Presentation Styles in Six Conditions (German and Japanese Subjects), SDs in Parentheses*

	Haptic		Color		Congruence		Incongruence		Blindfolded		Looking	
	t	pref	t	pref	t	pref	t	pref	t	pref	t	pref
German subjects	**2.33**	**0.83**	**2.70**	**-0.04**	**2.32**	**1.48**	**2.71**	**0.14**	**3.10**	**n/a**	**1.43**	**n/a**
	(0.44)	**(0.76)**	**(0.71)**	**(0.70)**	**(0.49)**	**(0.39)**	**(0.75)**	**(1.01)**	**(0.87)**		**(0.36)**	
BIG-SMALL metaphors	1.86	0.93	2.15	-0.32	1.88	1.29	2.10	0.46	2.28	n/a	1.48	n/a
	(0.50)	(1.21)	(0.73)	(1.02)	(0.49)	(1.01)	(0.66)	(1.32)	(0.70)		(0.58)	
HEAVY-LIGHT metaphors	2.06	0.86	2.74	-0.04	2.11	1.71	2.51	-0.04	2.98	n/a	1.34	n/a
	(0.51)	(0.97)	(0.79)	(1.10)	(0.58)	(0.46)	(0.73)	(1.29)	(0.89)		(0.36)	
STRONG-WEAK metaphors	2.81	0.96	3.18	0.04	2.79	1.46	3.13	0.07	4.09	n/a	1.47	n/a
	(0.61)	(1.07)	(1.01)	(1.00)	(0.76)	(0.69)	(0.89)	(1.21)	(1.51)		(0.44)	
WARM-COLD metaphors	2.57	0.57	2.74	0.14	2.48	1.46	3.10	0.07	3.06	n/a	1.43	n/a
	(0.52)	(1.00)	(0.66)	(0.89)	(0.53)	(0.92)	(0.99)	(1.36)	(0.87)		(0.38)	
Japanese subjects	**2.90**	**0.61**	**3.12**	**0.32**	**2.88**	**1.16**	**3.33**	**-0.17**	**3.77**	**n/a**	**1.74**	**n/a**
	(0.52)	**(0.61)**	**(0.66)**	**(0.50)**	**(0.57)**	**(0.61)**	**(0.77)**	**(0.61)**	**(0.80)**		**(0.68)**	
BIG-SMALL metaphors	2.75	0.69	2.72	0.62	2.77	0.78	2.95	0.24	3.15	n/a	1.78	n/a
	(0.53)	(1.02)	(0.64)	(0.96)	(0.60)	(1.22)	(0.75)	(1.07)	(0.78)		(0.69)	
HEAVY-LIGHT metaphors	2.83	0.67	3.17	0.38	2.75	1.31	3.19	-0.47	3.54	n/a	1.74	n/a
	(0.53)	(1.09)	(0.75)	(0.94)	(0.62)	(1.10)	(0.79)	(0.94)	(0.87)		(0.67)	
STRONG-WEAK metaphors	3.01	0.62	3.46	0.09	3.02	0.91	3.49	0.04	4.74	n/a	1.76	n/a
	(0.59)	(1.13)	(0.90)	(0.97)	(0.65)	(1.20)	(0.92)	(1.04)	(1.25)		(0.82)	
WARM-COLD metaphors	3.01	0.44	3.15	0.20	2.97	1.62	3.67	-0.49	3.64	n/a	1.68	n/a
	(0.59)	(0.99)	(0.58)	(0.92)	(0.57)	(0.83)	(0.95)	(0.87)	(0.91)		(0.75)	
Overall	**2.65**	**0.69**	**2.95**	**0.18**	**2.64**	**1.28**	**3.07**	**-0.05**	**3.49**	**n/a**	**1.61**	**n/a**
	(0.56)	**(0.68)**	**(0.71)**	**(0.61)**	**(0.60)**	**(0.55)**	**(0.81)**	**(0.80)**	**(0.89)**		**(0.59)**	

In order to get answers to hypotheses a-d), multiple repeated measures *t*-tests and MANOVAs were computed. Note that instead of applying Bonferroni-corrections to account for alpha-error accumulation, exact *p*-values and effect sizes are reported (Nakagawa, 2004). The reported degrees of freedom may differ between different conditions because of missing data due to technical problems while recording reaction times. For hypotheses a-d) the data of German and Japanese participants were averaged. The results are grouped according to the four hypotheses.

(a) To test the effect of *congruency* and its impact on the effectiveness and efficiency of the mapping, the *haptic-color congruence* and *haptic-color incongruence* condition were compared regarding matching performance and times for each image schema and across all four using multiple *t*-tests, see Tab. 6.13. For all image schemas, the metaphor consistency in the answers (effectiveness) was higher in the *haptic-color congruence* compared to the *haptic-color incongruence* condition. Subjects matched objects in the *haptic-color congruence* condition faster compared to the *haptic-color incongruence* condition.

Table 6.6.: *Results of Single Comparisons Between Congruent and Incongruent Mappings Regarding Effectiveness (%) and Efficiency (t in Seconds) Using Paired t-tests*

		Haptic-color congruence		Haptic-color incongruence					
Parameter	Image Schema	*M*	*SD*	*M*	*SD*	*t*	*df*	*p*	*d*
%	BIG-SMALL**	89	17	65	31	5.67	71	.003	0.96
	HEAVY-LIGHT***	96	9	70	23	8.85	71	.000	1.49
	STRONG-WEAK***	92	12	74	24	6.07	71	.000	0.95
	WARM-COLD***	89	12	67	31	5.73	71	.000	0.94
	Overall* **	**91**	**13**	**63**	**27**	**10.71**	**71**	**.000**	**1.32**
t	BIG-SMALL***	2.40	0.71	2.59	0.83	3.10	63	.000	0.25
	HEAVY-LIGHT***	2.48	0.67	2.91	0.83	5.90	63	.000	0.57
	STRONG-WEAK***	2.92	0.70	3.34	0.91	5.23	63	.000	0.52
	WARM-COLD***	2.77	0.60	3.43	1.00	7.82	63	.000	0.80
	Overall* **	**2.64**	**0.60**	**3.07**	**0.81**	**8.24**	**63**	**.000**	**0.60**

The second part of the hypothesis stated that participants would prefer metaphorically congruent objects over metaphorically incongruent objects, no matter if the image-schematic source dimension is instantiated through haptic properties, color, or both. Therefore, the data from the preference questionnaire of the *haptic*, *color* and *haptic-color congruence* conditions were compared to the *haptic-color incongruence* condition. In Tab. 6.7 the results of multiple t-tests are shown. Overall, the *haptic* and *haptic-color congruence* condition were always preferred compared to the *haptic-color incongruence* condition. The *color* condition was only preferred over the *haptic-color incongruence* condition for the image schemas HEAVY-LIGHT and WARM-COLD.

Table 6.7.: *Results of Single Comparisons Between Different Image Schema Instantiations Regarding Preference Ratings Using Paired t-tests*

Image schema	Condition	$t(72)$	p	d
BIG-SMALL	haptic-incongruent**	2.71	.008	0.40
	color-incongruent	0.37	.713	
	congruent-incongruent**	3.16	.002	0.55
HEAVY-LIGHT	haptic-incongruent***	5.55	.000	1.07
	color-incongruent*	2.50	.015	0.58
	congruent-incongruent***	9.63	.000	1.84
STRONG-WEAK	haptic-incongruent***	3.70	.000	0.63
	color-incongruent	0.08	.941	
	congruent-incongruent***	5.12	.000	0.99
WARM-COLD	haptic-incongruent***	4.50	.000	0.72
	color-incongruent*	2.47	.016	0.45
	congruent-incongruent***	10.21	.000	1.83
Overall	**haptic-incongruent*** **	**6.20**	**.000**	**1.00**
	color-incongruent	**1.75**	**.084**	
	congruent-incongruent* **	**10.73**	**.000**	**1.93**

(b) In order to investigate the second hypothesis regarding the influence of the interaction modality, the effectiveness and efficiency of the mapping were compared when participants investigated the objects by vision and touch or by only one of the modalities. For BIG-SMALL and WARM-COLD, the same pattern emerged: while the consistency in the answers did not differ when participants lifted objects while looking at them or being blindfolded, their performance was higher when they touched and looked at the objects compared to when they were not allowed to touch them. In the HEAVY-LIGHT condition, neither the *haptic* vs. *blindfolded* nor *color* vs. *only looking* performance differed significantly from each other. For

STRONG-WEAK, participants had more matches with the image-schematic metaphor in the *haptic* condition compared to the *blindfolded* condition. As for the *color* vs. *only looking* conditions, participants performed equally well. Across all four image schemas, participants matched the objects to the metaphor with a higher consistency in the *haptic* (compared to the *blindfolded*) and *color* (compared to the *only looking*) condition. Regarding the time required for the task, the data shows a clear pattern. For all image schemas and across all four of them, participants made their responses faster when they could inspect the objects via touch and vision compared to when they were blindfolded. They were also faster when they were not allowed to touch colored objects compared to looking at them while lifting. The detailed results of the analyses are depicted in Tab. 6.8.

Table 6.8.: *Results of Single Comparisons Between Different Interaction Modalities Regarding Effectiveness (%) and Efficiency (t in Seconds) Using Paired t-tests*

		Haptic		Blindfolded					
Parameter	Image Schema	*M*	*SD*	*M*	*SD*	*t*	*df*	*p*	*d*
%	BIG-SMALL	85	22	85	22	0.09	71	.930	
	HEAVY-LIGHT	94	10	91	14	1.75	71	.084	
	STRONG-WEAK*	91	14	84	21	2.50	71	.015	0.39
	WARM-COLD	88	12	86	14	1.30	71	.197	
	Overall**	**90**	**15**	**87**	**18**	**2.73**	**71**	**.008**	**0.36**
t	BIG-SMALL***	2.37	0.68	2.78	0.86	4.87	62	.000	0.52
	HEAVY-LIGHT***	2.51	0.64	3.31	0.91	8.42	63	.000	0.94
	STRONG-WEAK***	2.93	0.60	4.47	1.39	9.98	63	.000	1.44
	WARM-COLD***	2.82	0.60	3.40	0.93	5.41	63	.000	0.72
	Overall***	**2.65**	**0.56**	**2.78**	**0.86**	**9.85**	**63**	**.000**	**1.12**
		Color		Only Looking					
Parameter	Image Schema	*M*	*SD*	*M*	*SD*	*t*	*df*	*p*	*d*
%	BIG-SMALL**	86	22	80	22	2.91	71	.005	0.27
	HEAVY-LIGHT	84	22	84	21	0.04	71	.966	
	STRONG-WEAK	86	21	88	17	0.59	71	.556	
	WARM-COLD**	88	13	83	11	2.81	71	.006	0.42
	Overall*	**86**	**20**	**83**	**18**	**2.08**	**71**	**.042**	**0.23**
t	BIG-SMALL***	2.48	0.73	1.65	0.66	12.17	63	.000	1.19
	HEAVY-LIGHT***	2.99	0.79	1.57	0.59	16.20	63	.000	2.04
	STRONG-WEAK***	3.34	0.95	1.63	0.69	16.69	63	.000	2.05
	WARM-COLD***	2.97	0.64	1.57	0.63	17.40	63	.000	2.20
	Overall***	**2.95**	**0.71**	**1.61**	**0.59**	**19.61**	**63**	**.000**	**2.05**

(c) The third hypothesis proposed that the effectiveness of a multimodal instantiation of an image schema is higher than a haptic instantiation, which is in turn more reliably identified as a color instantiation. The results of the multiple comparisons are shown in Tab. 6.9. Overall, the effectiveness of the mapping did not differ between uni- or multimodal instantiations, as long as the haptic domain was involved. Metaphoric extensions of HEAVY-LIGHT and STRONG-WEAK (but not BIG-SMALL and WARM-COLD) were better identified in the multimodal instantiation compared to color, leading to an overall advantage of the *haptic-color congruence* condition compared to the *color* condition. Finally, the haptic instantiation of the image schema only outperformed the color instantiation for the HEAVY-LIGHT image schema. This effect was so strong that the comparison between the *haptic* and *color* condition reached significance in favor of the haptic instantiation.

Table 6.9.: *Results of Single Comparisons Between Different Image Schema Instantiations Regarding Effectiveness (%) Using Paired t-tests*

Image schema	Condition	$t(71)$	p	d
BIG-SMALL	congruent > color	1.05	.295	
	haptic > color	0.18	.859	
	congruent > haptic	1.04	.302	
HEAVY-LIGHT	congruent > color***	4.71	.000	0.71
	haptic > color***	3.74	.000	0.59
	congruent > haptic	1.29	.203	
STRONG-WEAK	congruent > color*	2.55	.013	0.35
	haptic > color	1.64	.106	
	congruent > haptic	0.57	.572	
WARM-COLD	congruent > color	1.11	.269	
	haptic > color	0.19	.847	
	congruent > haptic	0.92	.358	
Overall	**congruent > color*** **	**5.06**	**.000**	**0.65**
	haptic > color **	**2.78**	**.007**	**0.43**
	congruent > haptic	**1.08**	**.288**	

The second part of the third hypothesis states that choices in the *color* condition should differ from chance, as colors indicate image schemas and therefore bias judgements in domains linked by metaphor. To test this hypothesis, the average matching performance in the *color* condition was compared to choices made by chance (50%) for each image schema and across all four image schemas. A one-samples *t*-test revealed that all *color* condition differ significantly from chance, see Tab. 6.10.

Table 6.10.: *Results of Single Comparisons of the Effectiveness (in%) of Color Instantiations Against Agreement by Chance Using One-Sample t-tests*

Image Schema	*M*	*SD*	*t(71)*	*p*	*d*
BIG-SMALL***	86	22	13.83	.000	0.50
HEAVY-LIGHT***	84	22	13.31	.000	0.47
STRONG-WEAK***	86	21	14.41	.000	0.50
WARM-COLD***	88	13	23.79	.000	0.53
Overall* **	**86**	**20**	**27.92**	**.000**	**0.51**

(d) In order to answer if the German and Japanese subjects performed differently, their overall responses (metaphor consistency in answers, reaction times, preference ratings if applicable) for the six different conditions were compared using independent measures MANOVA and the results are shown in Tab. 6.11. As for matching performance, Japanese subjects had different values in two out of six conditions. In terms of reaction times, Japanese subjects needed significantly more time before they subjected their answers in all conditions. Regarding the preference data, the Japanese subjects rated two out of four conditions differently than the German subjects.

Table 6.11.: *Results of the MANOVAs Between German and Japanese Subjects Regarding Effectiveness (%), Efficiency (t in Seconds) and Preference Ratings (-2 to +2)*

Parameter	Condition	Result	F	df	p	η_p^2
%	haptic	Germans = Japanese	1,70	0.12	.732	
	color	Germans = Japanese	1,70	3.88	.053	
	haptic-color congruence	Germans = Japanese	1,70	0.34	.559	
	haptic-color incongruence	Germans < Japanese***	1,70	17.14	.000	.197
	blindfolded	Germans = Japanese	1,70	1.25	.267	
	only looking	Germans > Japanese**	1,70	10.64	.002	.132
t	haptic	Germans < Japanese***	1,61	21.13	.000	.257
	color	Germans < Japanese*	1,61	5.88	.018	.088
	haptic-color congruence	Germans < Japanese***	1,61	16.19	.000	.210
	haptic-color incongruence	Germans < Japanese**	1,61	9.93	.003	.140
	blindfolded	Germans < Japanese**	1,61	10.14	.002	.143
	only looking	Germans < Japanese*	1,61	4.28	.043	.066
pref	haptic	Germans = Japanese	1,71	1.94	.168	
	color	Germans < Japanese*	1,71	6.78	.011	.087
	haptic-color congruence	Germans > Japanese *	1,71	6.45	.013	.083
	haptic-color incongruence	Germans = Japanese	1,71	2.67	.107	

The second part of the fourth hypothesis concerned the language dependency of a conceptual metaphor. Three metaphors could not be fully identified in Japanese, although they are present in German language: DIFFICULT IS HEAVY - EASY IS LIGHT, MALE IS STRONG - FEMALE IS WEAK, and PROBLEMATIC IS WARM - UNPROBLEMATIC IS COLD. It was expected that Japanese could still reliably identify this mapping similar to the German participants. Independent-sample t-tests were performed between the efficiency of the mapping of German and Japanese subjects for the *haptic, color, haptic-color congruence* and *haptic-color incongruence* condition, and the results are presented in Tab. 6.12. Three comparisons were significant in a way that the mapping worked even better for the Japanese than for the German subjects. For the other comparisons, no difference in the consistency in the participants' answers could be found between both cultural groups.

Table 6.12.: *Results of Single Comparisons of Three Metaphors Between German and Japanese Subjects Regarding Effectiveness (in %) Using Independent t-tests*

Metaphor	Condition	*t*	*df*	*p*	*d*
DIFFICULT IS HEAVY -	haptic	0.32	70	.748	
EASY IS LIGHT	color** (Germans < Japanese)	2.93	30,726	.001	0.78
	haptic-color congruence	1.49	70	.429	
	haptic-color incongruence	1.58	67,274	.120	
MALE IS STRONG -	haptic* (Germans < Japanese)	2.33	31,522	.026	0.60
FEMALE IS WEAK	color	0.46	70	.645	
	haptic-color congruence	0.89	38,480	.377	
	haptic-color incongruence* (Germans < Japanese)	2.164	70	.034	0.52
PROBLEMATIC IS WARM -	haptic	0.81	70	.421	
UNPROBLEMATIC IS COLD	color	0.52	57	.605	
	haptic-color congruence	1.46	61,451	.149	
	haptic-color incongruence	0.65	55	.518	

The results of the four hypotheses and accompanying predictions tested in Study 5 are summarized in Tab. 6.13.

Table 6.13.: *Summary of the Hypotheses and Results of Study 5*

	Matching Consistency	Time	Preference
(a)	congruent > incongruent ✓	congruent < incongruent ✓	haptic > incongruent ✓
			color > incongruent (✓)
			congruent > incongruent ✓
(b)	haptic > blindfolded ✓	haptic < blindfolded ✓	
	color > only looking ✓	only looking < color ✓	
(c)	congruent > color ✓		
	haptic > color ✓		
	congruent > haptic ✗		
	color ≠ chance ✓		
(d)	Japanese = German (✓)	Japanese = German ✗	Japanese = German (✓)
	Japanese = German for Metaphor $_{\text{only in German}}$✓		

Note. Results are marked with ✓ if the hypothesis was confirmed, and ✗if not confirmed. Results that only partly support a hypothesis are indicated by (✓).

After the experimental trials were conducted, participants were asked how they made their choices in an open answer style. Overall, 75% of the subjects mentioned that their decisions were based on the haptic attributes of the objects, but also on color (named by 75%). 36% explained that they made their choices intuitively, following their gut feelings. 7% said that some adjectives represented synonyms of the haptic properties of the objects and their decisions were based on such parallels (for example, one participant said that he or she *feels heavy when being sad* and therefore made a link; another subject mentioned that some words can be used as synonyms like *heavy* for *difficult*). There were no differences between the answers of the German and Japanese subjects.

6.1.4. Discussion

6.1.4.1. General Discussion and Underlying Mechanisms

The majority of the empirical results of Study 5 are in line with the predictions made by the four hypotheses derived from CMToC. First, the overall results, followed by the single comparisons for each image schema and condition are discussed here.

According to the overall test, it is concluded that by determining cross-modal associations between colors and other physical properties from the literature, it could be predicted how colors represent abstract concepts that are perceptually grounded in experiences with temperature, size, strength and weight. When an image schema was instantiated through an indicating color, participants frequently assigned colored objects to corresponding metaphorical extensions of those image schemas. As tested with the second part of the hypothesis c), the majority of the subjects of Study 5 (86%) chose the object that best represented a metaphorically related abstract concept, differing significantly from agreement by chance for all image schemas. Since color is only one indicator of object and material identity that grounds abstract concepts, the prediction that instantiations that include a haptic property will outperform pure color instantiations received overall support with medium effect sizes (d = 0.43 for *haptic* > *color*, d = 0.65 for *haptic-color congruence* > *color*), as tested with the first part of hypothesis c). Regarding the individual image schemas, the conditions did not differ at all for BIG-SMALL and WARM-COLD, and the *haptic* and *color* condition did not differ for STRONG-WEAK. A difference could be observed for the *haptic-color congruence* and *color* condition in the STRONG-WEAK image schema, as well as for both comparisons in HEAVY-LIGHT. This points to a weaker influence of color properties in metaphorical dimensions of mechanical stimuli like weight and force, and a higher influence in metaphorical extensions of visual and tactile properties like size and temperature. This was not expected from the regression models determined in Study 1, which showed that the visual/tactile stimuli size and temperature predict 8% less of the variation in color ratings compared to the mechanical stimuli weight and force ($R^2_{\text{mechanical stimuli}}$ = .900, $R^2_{\text{visual/tactile stimuli}}$ = .824).

Since the *haptic-color congruence* and *haptic* conditions always received a higher agreement than the *color* conditions, instantiating the image schema haptically elicits the strongest association with metaphorically linked concepts. However, conveying abstract content only through color is still a viable approach that only resulted in a decrease in mapping effectiveness of 4% compared to the *haptic* instantiation and 5% compared to the *haptic-color congruent* instantiation in this experimental setup. Therefore, following CMToC does not only shed light on how haptic attributes like temperature, size, strength and weight are best matched with color, but also on how color might facilitate abstract thought.

The second part of hypothesis a) aimed at comparing instantiation preferences for different metaphors. Each participant rated pictograms of all four instantiations regarding one metaphorical dimension. Clearly, *congruent* instantiations in which both color and haptic property conveyed the same image schema were most preferred, especially compared to *incongruent* instantiations. Overall and for each individual schema, objects in the *haptic* condition received also higher preference ratings compared to the *incongruent* pair. Only the *color* condition was not (BIG-SMALL and STRONG-WEAK) or only slightly (HEAVY-LIGHT and WARM-COLD) preferred over the *incongruent* condition. This means that although color was frequently matched to the linked image-schematic metaphor, the participants opted for representations that include a haptic instantiation of the image schema. Therefore, the preference ratings are in line with the effectiveness and efficiency data. One reason why the *color* instantiation was less preferred could be the general ambiguity of color signals, which get contextualized through haptic properties. For example, the colors used to represent weight, size and strength can be used interchangeably, since all dimensions depend mainly on brightness and saturation.

In general, following the results gathered for hypothesis a), it is recommended that color should be either aligned with haptic properties (*haptic-color congruence* condition) or a neutral color should be chosen with regard to the haptic property that is in focus of the interaction (*haptic* condition), thereby eliminating an ambiguous information channel. The greatest performance loss both in effectiveness and efficiency could be observed in the *haptic-color incongruence* condition compared to the *haptic-color congruence* condition (-28% metaphor consistency in the answers, +17% more time). This loss in performance was consistent over all image schemas. The large congruity effect obtained between *congruent* and *incongruent* trials implies that the participants were sensitive to the congruence or incongruence of the perceptual whole. Following the time they needed to make their choices and the feedback they provided, it is unlikely that the congruency is located at the perceptual stage of stimulus processing. A post-perceptual model would explain that different competing response tendencies exist (Melara, 1989). While the haptic property invokes one response tendency, color invokes another. If both tendencies are the same (*haptic-color congruence* condition), a congruency effect results. If both tendencies are different (*haptic-color incongruence* condition), interference is the consequence. Semantic involvement as an alternative explanation can also account for the congruity effect. According to a semantic explanation, the co-usage of attribute image schemas for haptic properties, abstract concepts as well as to describe color properties may be the basis for the figurative mappings obtained in Study 5. For example, the word *heavy* is used to describe darker and more saturated colors, weight, as well as *difficulty*. Following a semantic model, the physical attributes (haptic, color) are first recorded into their corresponding discrete tags (Banks & Root, 1979; Melara, 1989), and interact because of their semantic similarity or dissimilarity. This interaction can take several forms. For instance, the darker color of a light box (weight task, *haptic-color incongruence* condition) activates a memory location that is semantically distant from the location activated by the low weight. The participant then has to decide which attribute and corresponding memory location is more critical to fulfill the task, and any misallocation of attention to the other memory location would result in a slowed classification (Dykes & Cooper, 1978). In a more indirect semantic account, similar instantiations of an image schema (e.g., a dark heavy box) could activate pathways that lead to the same response, whereas incongruent attributes (e.g., a bright heavy box) activate different processing pathways and cause interference (Posner & Snyder, 1975). Here, the mapping of semantic content to response tendencies would happen at a more prior stage of information processing (Melara, 1989). Independently if the

the processes are more direct or indirect, semantic processes might be the locus of the observed congruency effects in Study 5.

In 2012, Lakens came up with an alternative explanation for metaphor congruency effects in general. In most studies to date, metaphor congruency effects were interpreted as response interferences (Lakens, 2012). However, structural overlap between bipolar concepts might provide a more parsimonious explanation. Lakens describes the phenomenon as follows: "Dimensions consisting of polar oppositions (e.g., good–bad or UP-DOWN) have a default (polar) endpoint (i.e., good, UP) that receives a processing benefit compared to the opposite (polar) endpoint (i.e., negative, DOWN). The reaction times for bipolar stimuli in bimanual categorization tasks can be accurately predicted based on their polarity benefits" (Lakens 2012, p.2). The image schemas investigated in Study 5 together with their metaphorical extensions are all polar opposites (Hampe, 2005c). Usually, the linguistically unmarked dimension which is the name-giving category is +*polar*, for example significant (SIGNIFICANCE), happy (HAPPINESS), and active (ACTIVITY). +Polar endpoints tend to be used more frequently in language and also "have an intrinsic processing benefit compared to –polar endpoints" (Lakens 2012, p.2). The linguistically unmarked dimension (+*polar*) in the image schemas is named first: BIG-SMALL, HEAVY-LIGHT, STRONG-WEAK, WARM-COLD. According to the polarity correspondence account, participants should have matched the +*polar* instantiations of image schemas to +*polar* metaphorical endpoints faster compared to -*polar* instantiations to -*polar* metaphorical endpoints. In order to test this prediction, the overall response times for +*polar* words were compared to those of -*polar* words. A *t*-test for independent means with linguistic markedness as independent variable and time as dependent variable did not show any significant time difference between +polar words, $M = 2.75$, $SD = 0.71$, and -polar words, $M = 2.84$, $SD = 1.10$, $t(226) = 0.69$, $p = .492$. The speed with which participants matched metaphors was independent of their polarity to the objects that visually or haptically instantiated image schemas. However, it has to be noted that the subjects in Study 5 were not constrained by a time limit, which might have overwritten a possible polarity correspondence effect compared to a speeded choice task. Moreover, although the polarity of image schemas and metaphors can be easily determined, colors, with a few exceptions, have no polarity. Therefore, the polarity account can not explain the results of the *color* condition, in which the image schema was only instantiated through color and matched to adjectives representing abstract domains without resorting to CMT.

Hypothesis b) aimed at testing whether the participants' performance differed depending on interaction modality - when they were required to use hands and eyes to inspect the objects compared to only one modality. Regarding required time to indicate a decision, the subjects frequently matched objects to metaphors faster when they could touch and look (*haptic* condition) compared to when they were blindfolded, and faster when they were not allowed to touch the colored objects compared to being asked to look at them while lifting. In each condition, the experimenter carefully controlled for the start and end point of the time recording to avoid a systematic bias. For the conditions which involved vision and touch, time recording was started when the two objects together with the printed adjective were placed in front of the participants and they lifted the objects with both hands and finished when one object was placed back on the table. In the condition which did not involve vision, time was recorded when the participants lifted both objects and the experimenter read the adjective and stopped when one object was placed on the table. For the condition that did not involve touch, recording started when the tray was placed in front of the participants and stopped when they pointed to the chosen object. In addition, the influence of color was kept constant for the objects used in the *haptic* condition, as gray has no strong association with any of the investigated image schema dimensions. Across Japanese and German subjects, the reaction times were in the following order: *only looking* < *haptic* < *color* < *blindfolded*. Haptic sensations in a sense of nocioception or reflex arc-based reactions are said to be the bodies fastest reactions to sensory input because they do not require top-down processing before a reaction is formed, like it is the case in vision. In this experiment, however, subjects had to explore stimuli that were not harmful to them (*blindfolded* condition), therefore not triggering any fast reflexes in response to a source that might cause damage to the body. Hence, exploring the objects and their haptic differences by touch was much slower than identifying the objects by vision, as indicated by very large effect sizes. As participants were not allowed to touch the objects in the *only looking* condition, they only had to make a classification based on one stimulus dimension, color, which is faster than dividing attention to both haptic and color properties. Unsurprisingly, participants subjected the fastest responses in this condition - but also had a slightly lower number of consistent matches with the metaphor overall. This result suggests that a speed-accuracy tradeoff could have taken place: participants made the fastest responses when they only had visual input, but an additional haptic sensation increased consistency of the matching performance.

Having a look at the individual image schemas, BIG-SMALL and WARM-COLD as well as HEAVY-LIGHT and STRONG-WEAK performed more similar to each other. When attribute image schemas involving a physical force (strength, weight) were instantiated through color, the effectiveness of the mapping did not differ whether subjects touched the objects or not. Although not reaching significance in the HEAVY-LIGHT schema ($p = .084$), the participants in Study 5 had a slightly increased accuracy when they were allowed to look at and lift the gray objects compared to only lifting them. As for the image schemas that can be perceived through vision (size) and temperature perception (temperature), the effectiveness of the mapping was similar if they were operationalized in a haptic way, no matter if participants perceived them with one or two modalities. This is surprising for BIG-SMALL, as it can be experienced both visual and through touch. It would have been plausible if the participants showed a higher consistency in their answers in the *haptic* compared to the *haptic blindfolded* condition, because they perceived the image schema with two modalities. Regarding the trials in which size and temperature were operationalized through color, the effectiveness of the mapping was higher when participants used hands and eyes compared to only visually inspecting the objects, as the interaction was performed multimodal (Bertelson & Gelder, 2004). These different influences of image schematic color instantiations for those involving forces compared to those involving mainly visual and temperature cues could be an interesting topic for future research.

Finally, with the fourth hypothesis, it was investigated if the results of German and Japanese subjects differ considerably. Most notable, the Japanese participants submitted their responses consistently slower across all conditions compared to the German participants. Speculating, a reason for this slower decision making process might be inherent to the Japanese society, characterized by strict hierarchy, high uncertainty and risk avoidance as well as decision making by consensus (Hofstede, 2001). Regarding the effectiveness of the mappings, subjects of both cultures differed in two of six conditions (*incongruent* and *only looking)* with medium sized effects (J. Cohen, 1988)). Small differences could be found for the preference ratings of the *haptic-color congruence* and *color* condition. However, the majority of the data showed no differences between subjects of both cultural backgrounds, which is interpreted in favor of hypothesis d). The second part of the hypothesis concerned whether conceptual mappings also work in a culture in which no corresponding linguistic metaphor exists. Three metaphors were identified that were not surfaced in Japanese, but Japanese subjects compared to Germans still showed the same or better results in terms of effectiveness, efficiency and satisfaction. It is therefore concluded that the identified metaphoric mappings do work even if there is no counterpart in language.

The semantic color associations with abstract domains predicted by CMToC could not have been derived from other theoretical accounts in color psychology. Color-in-context theory predicts that people associate colors with those abstract concepts with which they were frequently co-experienced. However, the color substituted image-schematic metaphors as investigated in Study 5 are not necessarily originating from sensorimotor experience, and are instead a result of metaphorical transfer. Therefore, color-in-context theory would only predict those associations that stem from frequent experiential encounters. As it is difficult to determine which specific color associations are affected, the predictions are too vague to be of practical relevance.

Affective Meanings Systems makes the same prediction for abstract concepts as for image schemas: concepts with overlapping EPA-profiles will be associated. Metaphorical extensions of the image schemas BIG-SMALL, HEAVY-LIGHT, STRONG-WEAK and WARM-COLD were taken from Adams and Osgood (1973, p.324 f.) and combined with colors that share similar EPA-profiles (see Tab. 5.1). A dependency on hue, saturation and brightness is indicated by a change in the algebraic sign for a specific color attribute between the two metaphoric end points.

Table 6.14.: *Image-Schematic Extensions, Corresponding Hues and HSB Component Codings after Adams and Osgood (1973).*

Metaphor	Hue	H	S	B	Metaphor	Hue	H	S	B
ACTIVE	Red	+	+	0	INACTIVE	Yellow	+	-	0
BRIGHT (knowledgeable)	Blue/Green	-	+/-	0					
DISTANT	Yellow	+	-	0					
GOOD	Blue/Green	-	+/-	0	BAD	Yellow	+	-	0
GUILT	Yellow	+	-	0					
MALE	Blue/Green	-	+/-	0	FEMALE	White	0	0	+
POWERFUL	Red	+	+	0					
SAD	Yellow	+	-	0					

Note. +: red (hue)/ rich (saturation)/ bright (brightness); -: blue (hue)/ pale (saturation)/ dark (brightness); 0: neutral; n/a: no data available.

Due to incomplete data, only three opposing metaphors can be matched to color and the following predictions are derived: GOOD = blue/green - BAD = yellow and MALE = blue/green, ACTIVE = high saturation - INACTIVE = low saturation and low brightness - FEMALE = white and high brightness. The latter two were investigated in Study 5 and both were empirically supported as population stereotypes, as the majority of the participants (96% and 94%, respectively) chose the appropriately colored objects over the inappropriate one in the second condition (*color*).

The Ecological Valence Theory on Human Color Preferences predicts that the value which is assigned to a color corresponds to the average value assigned to the objects that carry this color. Although this theory cannot predict the outcomes of this experiment, it entails the assumption that people tend to prefer objects in their favorite colors are thus more likely to choose objects in their favorite color (hue). Therefore, despite other prevalent color associations with abstract (and often colorless) concepts, subjects would be biased towards their favorite color. As the participants' favorite colors in Study 5 were dominated by primary and secondary hues, the only suitable condition to test this prediction is the second condition (*color*) of the WARM-COLD image schema using blue and orange. In this condition, only the hue was varied between the shown objects (either *blue* or orange). If participants choose the object in their favorite color more often, it is expected that participants with *blue* as their favorite color (n = 14) will chose more often blue objects compared to participants with *orange* as their favorite color (n = 9). For none of the ten tested metaphors did the performance of both participant groups differ from each other ($p \geqq .125$). This means that despite different color preferences, the participants did not choose the object colored in their favorite hue more often than the other one, and were instead influenced by how well the color matched the image-schematic metaphor.

Overall, the empirical results in this experimental paradigm could be best predicted and explained by CMToC. Color-in-context theory could not make any precise predictions for the experimental paradigm of Study 5, and the prediction of Ecological Valence Theory on Human Color Preferences were not supported. Affective Meanings Systems could predict two of the 20 semantic color associations (10%) with abstract concepts, which were both well supported by the empirical results.

6.1.4.2. Experimental Design

One issue that might have played a role in Study 5 is the problem of a criterion bias. If answer categories contain a component that is part of the criterion under investigation, a systematic bias can result, affecting validity and reliability of the outcome (Brogden & Taylor, 1950). Basically, the metaphoric dimension in Study 5 was presented as linguistic stimuli, whereas the image schematic instantiations, as part of the metaphor in terms of the originating source concept, were instantiated physically. Presenting both metaphoric concepts and image schemas verbally might have otherwise resulted in a criterion bias, as image schemas function as linguistic synonyms of metaphoric content.

Moreover, the concrete instantiations of the image schemas might have played a role. As it has been shown in other studies investigating population stereotypes, the design of the experimental material greatly affected the results. For example, while Hurtienne et al. 2009 found support of BRIGHT-DARK mappings when the image schema was instantiated as black and white lego bricks, mappings did not work when the bricks were colored bright and dark green. Macaranas et al. found no support for FRONT-BACK mappings when the image schema was operationalized with a toy car taped on a cardboard ground with either the front or back half off the ground (Macaranas et al., 2012), whereas Hurtienne and Meschke found strong support for the mapping with an operationalization as labeled matchboxes showing feet with an arrow either pointing forward or backward (Hurtienne & Meschke, 2016). These examples illustrate that the consistency of participants' choices with the proposed metaphors vary as a function of the appropriateness of the material in representing the image schema. To ensure the appropriateness of the material, a manipulation check was performed before the image-schematic metaphors were tested in Study 5.

Another issue that needs to be mentioned is that the observed associations of the matching task in Study 5 can be an artifact of the method itself. In explicit forced choice tasks, especially those with two alternatives, it is possible that the participants only chose the option which was the least dissonant among otherwise unassociated concepts (Deroy & Valentin, 2011). The unlimited time the participants had may have reinforced this effect. One way to overcome this issue is to make use of more implicit measures such as the Implicit Association Test (Greenwald, McGhee, &

Schwartz, 2008) in which subjects have to make speeded classifications (Wan et al., 2014). Another possibility to reduce such an effect is to offer participants a no-choice option (Dhar & Simonson, 2003). In theory, such a response option draws proportionally from the available alternatives, so that the qualitative conclusion remains unaffected. In return, a no-choice response alternative competes with options that subjects would have chosen when they were uncertain. It therefore strengthens confident responses and decreases the relative share of an option that is average on all dimensions (Dhar & Simonson, 2003). Furthermore, according to Dhar and Simonson, a no-choice option reduces general discomfort of a forced choice task. Following these considerations, the effects in Study 5 tend to be overestimated as a result of the experimental paradigm.

6.1.4.3. Color Population Stereotypes

In Study 5, population stereotypes of visual and tangible object attributes were investigated. Population stereotypes occur when the population at large consistently chooses between physical and abstract options (Hurtienne et al., 2009). 20 physical-to-abstract mappings in the areas of temperature, size, strength and weight, were identified and validated with a total of 75 participants with either German or Japanese cultural background. Due to this limited and not representative sample size, only those metaphorical representations that were identified by the majority of the participants ($\geq 80\%$, $\varkappa \geq .60$) are recommended as viable population stereotypes and design guidelines. For example, the manipulation of physical weight, i.e., weight increase, predictably biased the categorization of objects toward "sad." Regarding color, the manipulation of an object's brightness predictably biased its categorization toward "sad", which is metaphorically linked to physical weight that is in turn indicated by an objects' brightness. While quantitative metaphoric concepts (more - less) received a very high agreement throughout all presentation styles and image schemas (i.e. they remained stable when metaphor-incongruent colors were applied), qualitative concepts like happiness were less reliably identified. This finding is in line with previous research on population stereotypes of visual and tangible object attributes (Hurtienne et al., 2009; Macaranas et al., 2012; Löffler et al., 2015, 2016). In addition, not all physical-to-abstract mappings work equally well and can be recommended. The metaphors KNOWING IS BIG - UNKNOWING IS SMALL and PROBLEMATIC IS WARM - UNPROBLEMATIC IS COLD were not empirically verified. Some qualitative feedback of the participants can shed light into why the KNOWLEDGE IS SIZE metaphor did not work well. One participant mentioned that *the small box is knowing, because it looks smart.* Another one said *knowledge is tightly packed in the small box, therefore it's knowing.* As for PROBLEMATIC IS WARM - UNPROBLEMATIC IS COLD, orange was not seen as a very alarming color. One participant said that if an organism reaches a very low temperature, this might cause problems, therefore matching COLD to PROBLEMATIC. Although the participants were instructed to choose objects as intuitively as possible, they did not receive any time limit and could freely associate, which may result in very subjective answers. It remains an open question for future research whether haptic related color properties can bias conceptual choices more subconsciously.

Another noteworthy aspect is that not all metaphorically congruent instantiations did perform equally well. In total, 48 out of 57 (84%) congruent representations (whether the image schema was operationalized through haptic or visual/color properties, or both) reached Kappa-values of at least 0.60, a typical finding in population stereotype studies based on image schemas (Hurtienne & Langdon, 2009; Macaranas et al., 2012; Löffler et al., 2015, 2016; Hurtienne & Meschke, 2016). Overall, color can function as a substitute for haptic properties, although some do perform decidedly worse (e.g. IMPORTANCE IS SIZE) and some better (e.g. HAPPINESS IS WARMTH) than a mere haptic instantiation, similar to what has been found in other experiments involving color-to-abstract mappings (Löffler et al., 2015, 2016). It is therefore argued that when a dynamic change in a physical property like temperature, size, strength or weight is hard to implement in a user interface, color-for-haptic-object substitution could be a viable approach. An open question for further research is if the influence of color on metaphorical extensions of force attribute image schemas (strength, weight) is different from the influence on temperature and size metaphors, as the results of Study 5 indicate.

A new aspect introduced in Study 5 are the split modality conditions in order to investigate each modalities' contribution to the overall performance (Schifferstein, 2006). Comparing the performance of perceiving image-schematic instantiations through one or two modalities, it was found that the matching speed to the metaphorically related concept was notably increased when participants only looked at colored objects. However, the matching accuracy was slightly decreased compared to perceiving objects via touch and vision. It is concluded that when the efficiency of the interaction is emphasized, the best results can be achieved by mapping metaphorical content to color without requiring the user to feel any haptic properties. In contrast, the greatest accuracy and highest preference ratings can be obtained when image schemas are represented through color and haptic properties (*haptic-color congruence* condition), since this condition did resolve perceptual ambiguities best (Lalanne & Lorenceau, 2004). Compared to that, the *haptic-color incongruence* condition, in which colors were counterintuitively applied to haptic properties, received the overall lowest accuracy and preference ratings and required participants the most time. This unveils the real power of making use of colors in user interfaces: incongruent colors can severely worsen the effectiveness, efficiency and obtained satisfaction of haptically instantiated metaphors. On the contrary, when color is aligned with the haptic instantiation, the physical-to-abstract mappings are perceived most intuitive and the ambiguity of whether a certain metaphor will be a population stereotype or not diminishes (17 out of 19 metaphors (90%) in the *haptic-color congruence* condition can be recommended as design guidelines).

According to the prior knowledge continuum of Hurtienne and Blessing, the sources can be organized from innate, to sensorimotor, to cultural and expertise knowledge

(Hurtienne & Blessing, 2007). As the image-schematic metaphors investigated in this study are motivated from sensorimotor experiences with the environment, which are rather invariable regarding the influence of cultural factors, it was expected that the answers participants give do not substantially vary among different cultures, even if no linguistic metaphor can be identified for one of the cultures. The performance and preferences between Germans and Japanese were compared, and they did not considerably differ. However, although showing the same response pattern, Japanese subjects were consistently slower compared to Germans. The results of this cross-cultural comparison demonstrate that by grounding color-to-abstract mappings in sensorimotor experience designers are able to use color in a way that is possibly valid across different languages and cultures, without having to rely on symbolic color meaning that is highly culturally dependent.

6.2. Chapter 6 Recap

Concluding, in Study 5, a broad scope of conceptual metaphors were decoupled from language discourse and transformed into visual and tangible objects and scenes that include different interaction aspects, for example the forcefulness of magnets. The results demonstrated that CMToC successfully predicts how abstract meaning can be conveyed through color in TUIs. Metaphorical extensions of image schemas are best identified when colors are used in an appropriate or neutral way with regard to the haptic property they indicate or correlate with. Because the color substituted image-schematic metaphors are based on sensorimotor knowledge of color relationships with physical properties, they are quite robust against cultural variation. Color can also be used as a substitute of sometimes difficult to implement or expensive haptic system output, which will increase efficiency if no haptic interaction is required, but also slightly decrease effectiveness and user satisfaction.

More studies are needed that identify further color substituted image-schematic metaphors, and that investigate these in more applied contexts than this initial study could provide. The experimental paradigm of two-alternative force choice tasks also has some drawbacks, like a general overestimation of the obtained effects. Although the participants' qualitative feedback in Study 5 showed that haptic properties and colors were consciously and equally considered when making their choices (mentioned by 75% of the participants), it remains an open question if color properties can bias metaphorically linked conceptual choices more subconsciously. This issue is addressed in the next chapter with a focus on metaphorical extensions of the image schema WARM-COLD.

7. Automatic Activation of Color-Temperature-Associations

First evidence shows that participants do consciously match colors to image schemas (Studies 1-4) and linked abstract domains (Study 5) in a way predicted by CMToC. To meet the precondition of intuitive use as subconscious application of prior knowledge, it has to be shown that these associations are also established automatically (Naumann et al., 2007). Moreover, the direction of semantic color associations is of potential interest: are the links between colors and physical properties (image schemas) as well as between colors and abstract domains bidirectional or only go in one direction? If the perception of color automatically influences the cognitive processing of associated concepts, designers can utilize color to support the semantic content they want to convey, even when users do not consciously think about the color design. For example, a website's dark background color can lead to a more negative affective interpretation of its content (Giron, 2016).

If it can be shown that the association also works in the other direction from cognitive processing of physical properties or abstract concepts to the perception of associated colors, designers have to consider that cognitive biases can lead to a processing benefit of specific colors or color distortion. This is relevant in situations when high color reproduction has to be achieved. For example, photographers and other professionals are concerned with translating screen content to a print with the highest color accuracy possible and therefore often use true color monitors; brand managers have to ensure a consistent brand image over various marketing channels, and when purchasing online, product photos have to show how the real product looks like as close as possible, otherwise resulting in customer dissatisfaction. In all these examples, the slightest distortion in color perception due to cognitive biases may result in profit loss and therefore should be avoided.

According to CMT, conceptual metaphors operate mostly unconscious and automatic (Lakoff, 1993) and "metaphorical mappings between dissimilar concepts tend to go in the direction of a concrete source concept to a relatively more abstract target concept, but not the other way around" (Landau et. al 2010, p. 1052). This cognitive asymmetry has been stressed by cognitive linguists to be a hallmark of metaphor (Lakoff & Turner, 2009; Landau, Meier, & Keefer, 2010; Ortony, 1979). Some scholars already found empirical support for this automaticity and asymmetry assumption in conceptual metaphors involving color. For example, Fetterman et al. investigated the conceptual metaphor ANGER IS SEEING RED, indicated in conventional speech such as "*seeing red*". In their 2012 paper, the implicit association between *anger* and *red* was addressed using a reaction time paradigm. While *red* font color facilitated participants' *anger* categorizations, *anger*-related words did not facilitate the categorization of *red* font color (Fetterman et al., 2012). However, many findings have been reported that demonstrate that the manipulation of an abstract concept leads to metaphor-consistent changes in the perception of more concrete domains (Landau et al., 2010), e.g., positive words are perceived as brighter (Meier et al., 2007). These mixed findings raise questions about whether semantic color associations automatically operate bidirectionally. The three experimental studies reported in this chapter aim to close this gap with regard to image schema-color metaphors and color substituted image-schematic metaphors by making use of a modified Stroop task.

The Stroop task has been used for more than half a century to study the implicit strength of associations between color and color-related words (McLeod, 1991). In the classic setup, subjects have to name colors of task-irrelevant semantically compatible and incompatible words. The resulting interference is calculated as the difference between times needed for the compatible vs. incompatible condition (called *Stroop Effect*), see Tab. 7.1.

Table 7.1.: *Conditions of a Stroop Task*

	Compatible Condition	Incompatible Condition
Stimulus	BLUE	BLUE
Response	blue	red
Interference	low	high

Years later, scholars modified the paradigm to study a variety of other associations. For example, Klein introduced a *semantic* variation of the Stroop task in which the relation of the presented words and colors was experimentally manipulated. The author could show that the more color-related the task-irrelevant word was, the higher the interference, opening the door for the study of a variety of other semantic color associations (Klein, 1964). For example, while the interference is highest in the standard Stroop condition (where color terms like *red* are used), interference is smaller but still robust for color-related words like *fire*. Other authors used the paradigm not only to investigate the effect of word content on color identification speed, but also the inverse, i.e., whether color perception impacts semantic identification speed, occasionally referred to as *Reverse Stroop Effect* (McLeod, 1991). An advantage of this paradigm over other implicit measures of associative strength like the *Implicit Association Test* or priming experiments is that both associative dimensions are integrated in a single stimulus, thereby avoiding effects of different presentation modes (e.g., color patches and linguistic stimuli).

The Stroop paradigm can be used to test the automaticity and asymmetry assumption in color associations with image schemas and abstract domains. Automaticity can be assumed when the task-irrelevant dimension (i.e., font color or semantic content) has an impact on the identification speed of the task-relevant dimension (i.e., semantic content or font color). Obtaining a Stroop Effect would indicate an association from semantic content (image-schematic or metaphoric) to color, while the Reverse Stroop Effect indicates an association from color to semantic content. The following three experiments test these predictions on the example of the WARM-COLD image schema and its metaphorical extensions. The well-documented WARM-COLD image schema was chosen as it plays an important role in social cognition (S. T. Fiske, Cuddy, & Glick, 2007) and scaffolds a variety of abstract concepts like *emotionality*, *psychological warmth*, and *activity*, that are potentially relevant for HCI, e.g. for computer-mediated communication or intimacy in distant relationships (Hassenzahl, Heidecker, Eckoldt, Diefenbach, & Hillmann, 2012).

7.1. Study 6 - Semantic Stroop Effect in Japanese Subjects

7.1.1. Hypotheses

Study 6 investigates whether presenting image schemas and metaphorical expressions (e.g. *cold* and *distant*, respectively) influences the identification speed of related colors (e.g., *blue*) in an automatic fashion (semantic Stroop Effect). The hypothesis of this experiment is that the stronger the association between a task-irrelevant but color-related word, the more interference is caused in color categorization (color > image schema > metaphor) relative to a non-word control condition. The experimental details, stimulus material and the results are described in the following paragraphs.

7.1.2. Method

7.1.2.1. Participants and general procedure

24 undergraduates (12 female, M_{age}= 20.8, SD_{age}= 1.36, range 18-23 years) were recruited from Nihon University in Japan and given course credit for their participation. All participants had normal or corrected-to-normal vision, were right-handed and spoke Japanese as their first language. Their favorite colors were blue (20.8%), red (16.7%), yellow (16.7%), orange (16.7%), black (8.3%), green (4.2%), violet (4.2%), and other (12.5%). After receiving general instructions and completing a demographic questionnaire, participants completed the color categorization task on a personal computer.

7.1.2.2. Stimuli

In total, 14 words were collected that represent color, the WARM-COLD image schema or a temperature-related image-schematic metaphor (see Tab. 7.2). Words indicating WARM and COLD were matched for length (number of Japanese characters), $t(10) = 0.71$, $p = .497$, syllables, $t(10) = 0.20$, $p = .845$, and word frequency, $t(10) = 0.31$, $p = .762$. Word frequency data was taken from the NTT Database Series, Lexical Properties of Japanese (Amano & Kondo, 1999). This word frequency is calculated from 14 years of Asahi Newspaper articles between 1985 and 1998 and indicates how many times a word appeared in this corpus. Each word was presented in 20 point Arial font, either in red (HSB = 30,94,98), blue (HSB = 240,66,88), green (HSB = 0,0,80) or violet (HSB = 289,91,88) on gray background (HSB = 0,0,80). The colors were matched in terms of WCAG2 foreground/background contrast (red 2.38, blue 3.77, green 2.17, violet 2.81).

Table 7.2.: *Overview of the Characteristics of the Stimulus Material Used in Study 6 (Translated from Japanese)*

Word	Condition	Temperature	Frequency	Length	Syllables
赤 (red)	color	warm	7552	1	2
青 (blue)	color	cold	4525	1	1
緑 (green)	color	neutral	6946	1	3
紫 (violet)	color	neutral	696	1	4
熱い (hot)	image schema	warm	4362	2	2
暖かい (warm)	image schema	warm	1373	3	3
冷たい (cold)	image schema	cold	3706	3	3
涼しい (cool)	image schema	cold	632	3	3
感情的な (emotional)	metaphor	warm	871	4	5
冷淡な (unemotional)	metaphor	cold	377	3	3
親密な (intimate)	metaphor	warm	786	3	4
よそよそしい (distant)	metaphor	cold	91	6	5
積極的な (active)	metaphor	warm	25482	4	7
消極的な (passive)	metaphor	cold	4757	4	6
XXXXX	control	neutral	-	-	-

Note. Word frequency shows how many times a word appeared in a selected corpus according to NTT Database Series, Lexical Properties of Japanese

7.1.2.3. Apparatus

The experiment was programmed and run using EPrime 2[1] on a 2.40 GHz computer with 4 GB RAM and a 60 Hz 15' LED color monitor. The monitor was placed about 50cm in front of the participants. The participants did use a common keyboard as input device. The keys *G, H, J* and *K* were marked with colored stickers according to the participants' response set.

7.1.2.4. Design and Procedure

Before working through the reaction time experiment, participants were asked to fill out a questionnaire to be completed with age, gender, current occupation, highest education, ethnicity, mother tongue and favorite color. In the main experiment, participants were instructed to categorize the font color of presented words as quickly and accurately as possible only using the index finger of their dominant hand (sec. A.6).

On each trial, a black fixation point appeared in the center of a gray screen for 300 ms. After that, a blank screen was shown for 500 ms which was replaced by the stimulus for a maximum duration of 3000 ms or until a response was given by the participant, see Fig. 7.1.

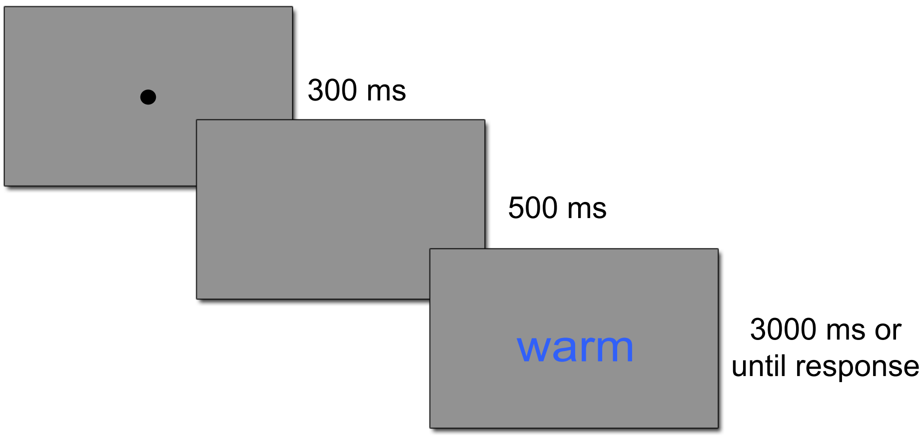

Figure 7.1.: Example of a trial display in Study 6.

The subjects were instructed to make as few errors as possible. The experimental trials were preceded by a practice block which showed each of the test stimuli once in random order. Only during the practice block, the participants received written feedback about their reaction times and whether their response was correct or not.

[1] www.pstnet.com

In the main part of the experiment, the participants had to work through eight blocks of stimuli grouped by condition (color, image schema, metaphor, control). To achieve an identical length per block despite the different numbers of stimuli in each condition, the stimuli were repeated until 48 were reached, ensuring that each stimulus was shown equally often per condition (least common multiple 12 times four colors). As each condition was repeated twice, the eight test blocks with 48 trials each made up a total of 384 trials per participant. After each block, the subjects were allowed to take a break. Within each block, the test trials were chosen randomly without repetition. All four colors were shown equally often (25%). While green and violet are unrelated to temperature, thus forming a neutral condition, red and blue were always shown incongruent to the word meaning (note that in this study, interference effects were compared to each other, not against congruent conditions). The whole procedure took about 30 minutes.

7.1.3. Results

Reaction times were handled in accordance with recommendations in the literature (M. D. Robinson, 2007). Inaccurate responses (2%) were deleted, reaction times were log-transformed, and trials which had reaction times 2.5 standard deviations below or above the grand mean of each condition were replaced with these values. In the image schema and metaphor condition, the trials were omitted in which green and violet were the font colors, as they are not temperature-relevant and were only included to keep the variance of the response set constant across all conditions. The average response time of the control condition in which only XXXXX was shown was subtracted from the response times of all other conditions (color, image schema, metaphor). This was done in order to eliminate the interference effect of the mere presentation of letters (Klein, 1964). As the three tested image-schematic color metaphors INTIMATE IS RED - DISTANT IS BLUE, EMOTIONAL IS RED - UNEMOTIONAL IS BLUE and ACTIVE IS RED - PASSIVE IS BLUE did not differ in their reaction times, $F<1$, they were averaged for the following analyses.

A repeated-measures ANOVA (color vs. image schema vs. metaphor) with a Greenhouse-Geisser correction was then conducted on the (trimmed) log reaction time means, though millisecond means and 95% confidence intervals (CIs) will be reported for ease of interpretation. Interference effects differed statistically significantly between conditions, $F(1.809,41.604) = 16.35$, $p < .001$, $\eta_p^2 = .415$. The interference effect of word content for font color categorization was smallest for metaphoric expressions (M = -2.62 ± 25.94 ms), followed by image schemas (M = 21.56 ± 30.92 ms) and colors (M = 46.64 ± 31.80 ms) (interference effects compared to the control condition). Colors were faster classified for metaphoric expressions compared to image schemas ($p = .006$), faster for image schemas compared to color words ($p < .001$), and faster for metaphors compared to colors ($p < .001$). Therefore, it can be concluded that the stronger the association between a task-irrelevant but color-related word, the stronger the elicited interference effect in color categorization.

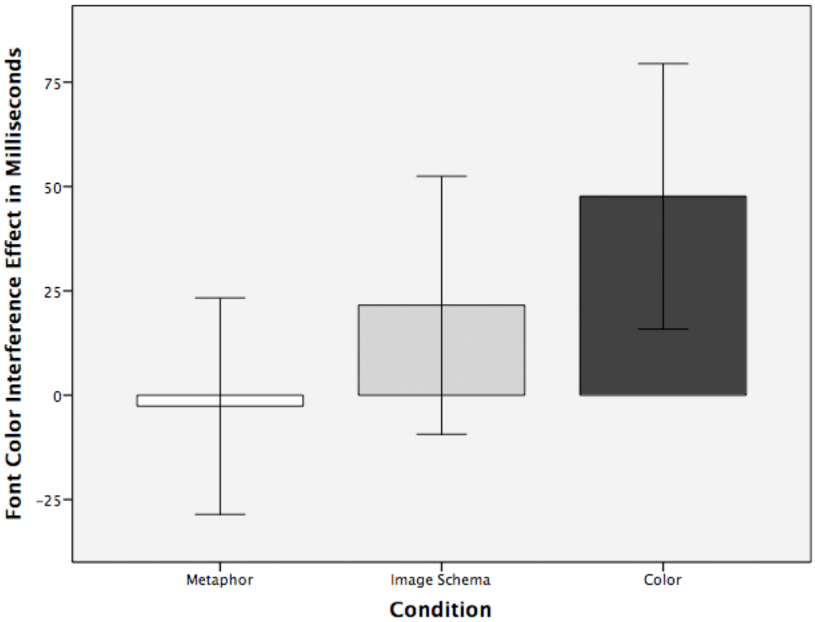

Figure 7.2.: Color categorization interference effect as a function of Condition (means with 95% CI error bars).

The length of the words, as well as their number of syllables and frequency in Japanese were controlled for *warm* compared to *cold*-related terms across all conditions. However, they differed between the conditions, as color names are very short and frequent in language compared to image schemas and metaphors. Shorter words with less syllables and words with higher frequency in language create larger interference effects independent of word content (McLeod, 1991). To investigate the impact of word length, number of syllables and word frequency on reaction times, a linear regression analysis was conducted with word length, number of syllables and word frequency as predictors. Overall, 7% of the variance in the reaction time data were explained by these three variables (adj. R^2= .07), F(3,68) = 2.77, p = .048, but no individual predictor reached statistical significance. A Pearson product-moment correlation revealed a significant relationship between word length and reaction times, r(72) = -.242, p = .040, as well as between syllables and reaction times, r(72) = -.243, p = .039, but not for reaction times and word frequency, r(72) = -.139, p = .246. The longer and the more syllables a word had, the faster the reaction times.

7.1.4. Discussion

Study 6 revealed a semantic Stroop effect: greater response speed interference for color categorization when the association between color and word was stronger. Metaphoric expressions were categorized faster compared to image schemas and color words. Color words, the classic Stroop condition, received the slowest overall response times. Thus, a semantic effect on color identification could be demonstrated, confirming the hypothesis.

A common explanation for these interference effects are that word-reading "is a fast, automatic, and unintentional process that recruits resources from the slower, less automatic task of color-naming" (Lorentz et al. 2016, p. 2). Such an impact of the WARM-COLD image schema as color-related words on color categorization has already been investigated in other studies using similar implicit measures. Ho et. al. applied an Implicit Association Task (IAT) as well as a priming study. In the IAT, the participants had to react to color patches (red, blue) and thermal words (warm, cold) with an either congruent or incongruent keyboard assignment. Facilitation effects could be obtained when red/warm and blue/cold were assigned in a congruent fashion compared to an incongruent assignment. Contrary, no influence of a color prime (red vs. blue) on the perception of an image-schematic target stimulus could be found (Ho, 2015). The authors' explanation is that the association between color and temperature is stronger in the direction of color to temperature. In addition, the stimulus onset asynchrony (SOA) between prime and target was 2000 ms - a long time interval that might have resolved prime processing before the target was perceived (Lorentz, Ekstrand, Gould, & Borowsky, 2016).

In a more similar experimental paradigm, Lorentz et al. investigated associations between *red* and synonyms of *hot* as well as between *blue* and synonyms of *cold* (Lorentz et al., 2016). One sub-task of the two conducted semantic Stroop experiments was to categorize font color. Interestingly, they only found a facilitation effect for color categorization performance when *red* font color matched *hot*-related concepts, but not for *blue* font color and *cold*-related concepts. Similar to Ho et al., the authors argue that this is because the association from temperature to color is less strong than the other way around, which was investigated with a semantic categorization task. In order to test whether *warm* and *cold* associations with color differ in strength, the log transformed average reaction times of the image schema condition in Study 6 (warm, hot vs. cold, cool) were subjected to a dependent-samples t-test. The result closely failed to reach statistical significance, $t(23) = 1.75$, $p = .093$. In tendency, the mean reaction time of the *warm* trials were indeed slightly slower ($M = 624.71$, $SD = 70.89$) than the reaction times of the *cold* trials ($M = 612.45$, $SD = 77.44$). It can be argued that this is due to a higher significance of *warm*-related associations for the human organism, but remains a topic to be explored in future research.

Contrary to image schemas, metaphoric expressions did not cause more interference than the non-word control condition. This result can be interpreted as that participants automatically established a link from the WARM-COLD image schema to related colors, but not from metaphorical extensions to color. Based on this finding, image schema-color metaphors fulfill the prerequisite for intuitive use, but color substituted image-schematic metaphors do not.

CMT predicts that metaphorical associations between two concepts tend to be stronger in the direction from (concrete or salient) *source* domain to (a more abstract and less salient) *target* domain. Image schema-color associations are associations between two concrete, physical domains. In metaphoric expressions, the image schema WARM-COLD appears as the *source* domain and color as *target* domain, e.g. *warm red*. According to this, the interference from semantic content to font color categorization is likely. Contrary to the experiential basis of the image schema-color association, temperature-related color substituted image-schematic metaphors like INTIMATE IS RED - DISTANT IS BLUE, EMOTIONAL IS RED - UNEMOTIONAL IS BLUE and ACTIVE IS RED - PASSIVE IS BLUE are not a result of frequent co-exposure of color and the specific abstract domain. Moreover, they do not result in linguistic metaphors and are, according to the results of Study 6, not activated automatically (direction abstract domain to color). This is a first hint that color substituted image-schematic metaphors may not qualify as conceptual metaphors that are subconsciously accessed.

However, Study 6 faces several shortcomings that limit the significance of the results. First, as the experiment did not involve a control for word length, number of syllables or word frequency between the three conditions, the interference effects are overestimated or might be even an artifact of the method. Although a regression analysis did not yield a substantial influence of these variables on reaction times, this alternative explanation cannot be fully ruled out. Second, *red* and *blue* in the response set (赤 akai - red, 青 aoi - blue) shared the same first letter like some of the image schemas and metaphors in the stimulus material, e.g. 熱い (atsui) - hot. This has been shown to bias results as well (McLeod, 1991). Because of these drawbacks and in order to test the robustness of the results, a conceptual replication of the experiment is reported in the next subchapter. The study design is complemented by a congruent condition as well as control words. In addition, Study 7 is conducted in Germany in order to test if similar results can be obtained with subjects of a different cultural background.

7.2. Study 7 - Semantic Stroop Effect in German Subjects

7.2.1. Hypothesis

Study 7 is a conceptual replication of Study 6 with German subjects. Changes were made in the stimulus and response set to avoid overlapping first letters. Moreover, a semantically congruent and neutral condition were introduced to be compared to the interference condition. As in Study 6, a semantic Stroop color-word interference task serves as the measure of implicit word-color association (McLeod, 1991) and it is expected that the interference caused is a function of the relationship between the word meaning and color (semantic gradient) (Klein, 1964). Color names (classical Stroop condition, e.g. *red*) and other color-related words (image-schematic (e.g. *cold*) and metaphoric expressions (e.g. *intimate*)) will be shown to the participants in different font colors (red, blue) which need to be categorized as fast as possible (semantic Stroop Effect). The hypothesis of this experiment is that the stronger the implicit association between a task-irrelevant but color-related word, the more interference is caused in color categorization (color > image schema > metaphor) relative to the congruent and neutral conditions. The experimental procedure, material and outcomes are described in the following.

7.2.2. Method

7.2.2.1. Participants and general procedure

The sample consisted of 30 undergraduate students from Würzburg University in Germany seeking course credit. All participants had normal or corrected-to-normal vision, and spoke German as their first language. Data of three participant was excluded because they were not right-handed. 27 subjects (16 female, M_{age}= 20.5, SD_{age}= 2.16, range 19-22 years) remained in the data set. Their favorite colors were blue (29.2%), green (25.0%), orange (12.5%), black (12.5%), red (8.3%), violet (4.2%), and other (4.2%). Identical to the procedure of Study 6, they received general instructions and then completed a demographic questionnaire, followed by a color categorization task on a personal computer.

7.2.2.2. Stimuli

In total, 12 test words were collected that represent color, the WARM-COLD image schema, a temperature-related image-schematic metaphor (see Tab. 7.3). For each of these test words, a word unrelated to *red* and *blue* hue as determined by a pre-test was selected and matched for length, number of syllables, word frequency, arousal and valence (control words). Word frequency as depicted in Tab. 7.3 was described as a function of frequency class (target word) = integer part ($\log_2$(frequency('der')/frequency(target word)))[2]. 'Der' is the most frequent word in German and serves as reference frame. For example, a frequency class of 11 means that 'der' is 2^{11} times more frequent than the target word. Arousal and valence values were adopted from (Hager & Hasselhorn, 1994), representing dimensional ratings from -3 to +3. No word in the stimulus set shared the first letter with the colors in the response set. Each word was presented in 20 point Arial font, either in red (HSB = 30,94,98) or blue (HSB = 240,66,88) on gray background (HSB = 0,0,80). Note that the response set was reduced to two WARM-COLD-relevant colors to increase the overall number of relevant trials. This response set size variation compared to Study 6 (four colors) is not likely to affect the experimental outcomes (McLeod, 1991).

[2]http://wortschatz.uni-leipzig.de/index.html, last accessed 18.09.2016

Table 7.3.: *Overview of the Characteristics of the Stimulus Material Used in Study 7 (Translated from German)*

Word	Condition	Temperature	Frequency	Length	Syllables	Arousal	Valence
rot (red)	color	warm	11	3	1	0.57	0.33
blau (blue)	color	cold	12	4	1	n/a	n/a
global (global)	control (color)	neutral	11	6	2	-0.38	-0.09
mobil (mobile)	control (color)	neutral	12	5	2	0.79	0.4
heiß (hot)	image schema	warm	10	4	1	n/a	n/a
warm (warm)	image schema	warm	11	4	1	-0.7	1.24
kalt (cold)	image schema	cold	10	4	1	-0.18	-1.06
kühl (cool)	image schema	cold	12	4	1	-0.21	-1.53
dumm (dumb)	control (image schema)	neutral	11	4	1	-0.72	-0.99
ehrlich (honest)	control (image schema)	neutral	10	7	2	-0.10	1.50
hart (hard)	control (image schema)	neutral	9	4	1	0.14	-1.53
hilfreich (helpful)	control (image schema)	neutral	11	8	2	1.43	1.80
emotional (emotional)	metaphor	warm	12	9	4	0.70	1.01
sachlich (objective)	metaphor	cold	12	8	2	-0.70	0.63
intim (intimate)	metaphor	warm	15	5	2	0.66	0.82
distanziert (distant)	metaphor	cold	13	11	3	-0.36	-0.40
aktiv (active)	metaphor	warm	9	5	2	1.46	1.12
passiv (passive)	metaphor	cold	13	6	2	-1.21	-1.08
ungenau (inaccurate)	control (metaphor)	neutral	14	7	3	0.12	-0.65
geschickt (skilled)	control (metaphor)	neutral	9	9	2	0.27	1.12
materiell (materially)	control (metaphor)	neutral	14	9	4	-0.32	-0.55
wach (awake)	control (metaphor)	neutral	12	4	1	0.46	0.72
virtuos (virtuoso)	control (metaphor)	neutral	14	7	3	1.36	1.33
höflich (polite)	control (metaphor)	neutral	13	7	2	-0.57	0.56
XXXXX	control	neutral	-	-	-	-	-

Note. The word frequency class shows how much more frequent the German word “der” is compared to the target word in the form $2^{\text{frequency class}}$ according to Wortschatz Portal Universität Leipzig

7.2.2.3. Apparatus

The experiment was programmed and run using EPrime 2[3] on a 3.40 GHz computer with 4 GB RAM and a 60 Hz FUJITSU B24W-6 LED color monitor. The monitor was placed about 50cm in front of the participants. The participants did use a common keyboard as input device. The keys *H* and *J* were marked with colored stickers according to the participants' response set.

7.2.2.4. Design and Procedure

The basic procedure was adopted from Study 6. Subjects first completed a demographic questionnaire, and went on with the main experiment. They were instructed to categorize the font color of presented words as quickly and accurately as possible only using the index finger of their dominant hand (see sec. A.7). The procedure of each trial was identical to Study 6: a black fixation point appeared in the center of a gray screen for 300 ms, followed by a blank screen for 500 ms and the stimulus for a maximum duration of 3000 ms or until a response was given by the participant. The preceding practice block contained 52 trials, which were taken from the stimulus set and presented in randomized order. Only in this practice block, the participants received written feedback about their reaction times and whether their response was correct or not. There were eight blocks of experimental trials consisting of 48 trials each, making a total of 384 trials per participant. Each condition (control, color, image schema, metaphor) was repeated twice. Subjects were allowed to take a break after each block. Within each block, the test trials were chosen randomly without repetition. Similar to Study 6, the items of each condition were repeated until a total number of 48 was reached within each block. *Blue* and *red* appeared equally often and were assigned to the color-related test words both in congruent and incongruent fashion. The whole procedure took about 30 minutes.

7.2.3. Results

Reaction times were handled identical to Study 6. First, inaccurate responses (3%) were deleted, reaction times were log-transformed, and trials which had reaction times 2.5 standard deviations below or above the grand mean of each condition (condition type/congruency/test vs. control) were replaced with these values. As the three tested color substituted image-schematic metaphors INTIMATE IS RED - DISTANT IS BLUE, EMOTIONAL IS RED - UNEMOTIONAL IS BLUE and ACTIVE IS RED - PASSIVE IS BLUE did not differ in their reaction times, $F<1$, they were averaged for the following analyses.

A 3 (Condition: color, image schema, metaphor) x 3 (Congruency: congruent, incongruent, neutral) repeated-measures ANOVA was conducted on the (trimmed)

[3]www.pstnet.com

log reaction time means, though millisecond means will be reported for ease of interpretation. There was a main effect for Condition, $F(2{,}68) = 16.32$, $p < .001$, $\eta_p^2 = .32$. Color words were categorized slower ($M = 467.83$, $SD = 66.60$) than metaphoric expressions ($M = 437.11$, $SD = 42.84$) or image schemas ($M = 434.82$, $SD = 42.50$). There was also a main effect for Congruency, $F(2{,}68) = 6.05$, $p = .008$, $\eta_p^2 = .15$. Incongruent trials had slower ($M = 456.59$, $SD = 59.80$) reaction times than congruent ($M = 444.03$, $SD = 47.40$) and neutral pairings ($M = 439.14$, $SD = 44.74$) of word content and color. Moreover, the Condition by Congruency interaction reached significance, $F(4{,}136) = 10.18$, $p < .001$, $\eta_p^2 = .23$. Millisecond means for this interaction are displayed in Fig. 7.3. The means reported in Fig. 7.3 suggest that font color categorization was impaired by semantically incongruent color word meaning, but not by congruent pairings or image-schematic and metaphorical expressions. This impression was confirmed by follow-up tests examining the effect of the congruency manipulation for each condition separately. Regarding the color condition, incongruent trials were categorized slower ($M = 494.92$, $SD = 91.52$) than congruent ($M = 452.27$, $SD = 55.81$) or neutral trials ($M = 456.30$, $SD = 52.47$). When subjects had to categorize font color of image-schematic or metaphorical content, the reaction times between congruent and incongruent trials did not differ from neutral words. The reaction times of all experimental conditions (color, image schema, metaphor) were higher than the control condition in which only XXXXX-strings were shown ($M = 422.91$, $SD = 53.45$), all $p < .050$. This is not surprising, as any semantic content will cause greater interference (Klein, 1964).

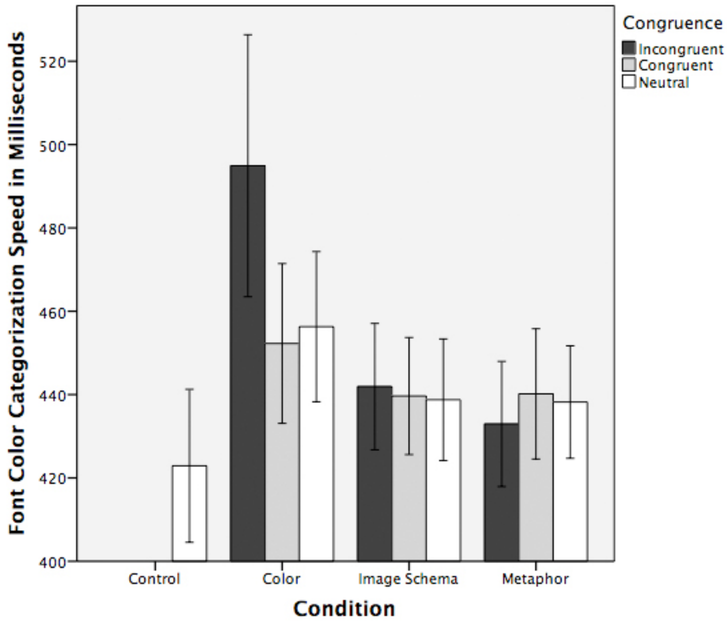

Figure 7.3.: Color categorization speed as a function of Condition and Congruency (means with 95% CI error bars).

Following the findings of Study 6 and the suggestions by Lorentz et al. 2016, possible differences in categorization performance between *red* and *blue* in experimental trails were examined. First, the conditions (color, image schema, metaphor), congruency types (congruent, incongruent, neutral) and font colors (red, blue) were visualized in a diagram, see Fig. 7.4. By visual inspection, the font color categorization performance does not seem to differ between red and blue font color, regardless of the condition and congruency type. Only the reaction times in the incongruent color condition seem to differ slightly. To test this impression, a dependent-samples t-test was conducted which confirmed that *blue* font color (M = 506.67, SD = 87.46) was categorized slower than *red* font color (M = 482.53, SD = 102.75) for incongruent trials in the color condition, $t(34) = 3.01$, $p = .005$, $d = 0.28$.

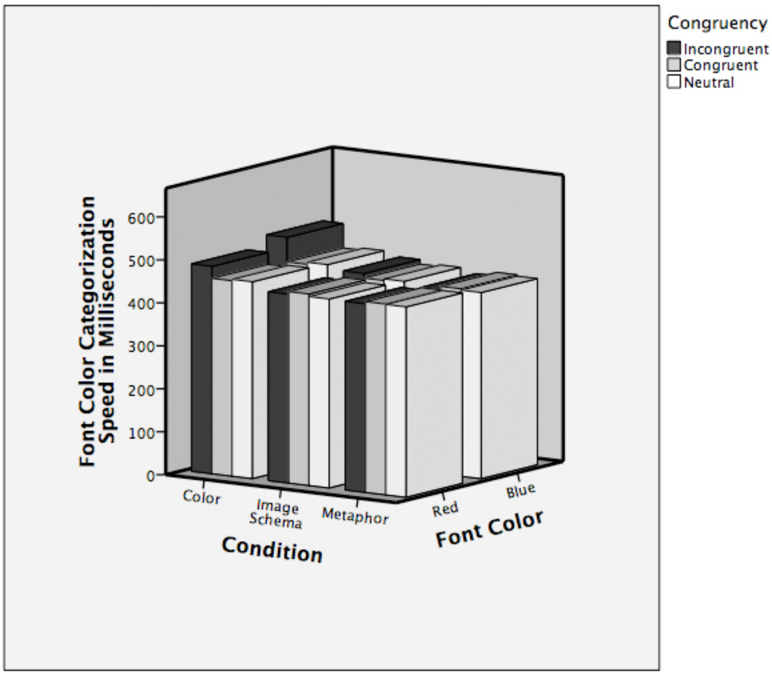

Figure 7.4.: Color categorization speed as a function of Condition, Congruency and font color.

7.2.4. Discussion

In Study 7, a robust Stroop effect could be obtained: response times for font color categorization were slowed when the word was presented incongruent to the font color (e.g., *red* written in *blue*). No indication of a facilitation effect of congruent trials compared to control words could be found. It was expected that if image schemas and abstract metaphoric concepts are automatically associated with colors, i.e., *red* and *blue*, processing these words will activate color representations and lead to interference and facilitation effects in font color categorization. This semantic Stroop effect occurs because not only color words but also words that denote color-diagnostic objects produce an interference in Stroop paradigms (Klein, 1964). It is likely that the interference caused by color-related words is due to an overlap in representational resources, which is supported by empirical findings that interference is also caused even if the color-related words are not linked to colors in the response set (Richter & Zwaan, 2009).

However, image-schematic or metaphoric semantic content had no impact on color categorization speed. According to this result, associations from image schemas and abstract domains to color are not activated automatically and therefore do not fulfill the prerequisite of intuitive use. The differences in response speed between the incongruent color, image schema and metaphor condition in Study 6 might therefore be attributable to the differences in word length and syllables between the conditions. One possible explanation why the semantic content of image-schematic and metaphoric expressions did not interfere with font color categorization is that the participants might not have processed the affective meaning of the words before categorizing them on font color (Lakens, 2012), which had just two response alternatives (for example, compared to the four response alternatives in Study 6). However, it is unclear why participants should have processed color words, but not image schemas and metaphors.

Another explanation might be that the associations between color and image schemas as well as between color and metaphors are less strong in the direction from semantic content to color than vice versa. Such an asymmetric relationship has already been found for example between anger and perceptual redness (Fetterman et al., 2012) as well as between valence and brightness (Meier et al., 2004). This is also in line with the less stable effect of semantic temperature words on font color categorization Lorentz et al. 2016 found in their recent study (Lorentz et al., 2016). Other studies, on the contrary, have revealed bidirectional associations between the (metaphoric) concept of *valence* and perceptual brightness (Meier et al., 2007; Okubo & Ishikawa, 2011; Banerjee, Chatterjee, & Sinha, 2012; Lakens et al., 2013). Thus, the conditions under which the relationship between color, image schemas and metaphoric dimensions is asymmetric remain a question for further research.

Study 8 aims at investigating the automatic relationship between color and image schemas as well as metaphors in the other direction: does perceiving different font colors have an influence on the semantic categorization of image-schematic and metaphoric content?

7.3. Study 8 - Reverse Semantic Stroop Effect in German Subjects

7.3.1. Hypothesis

To the extent that colors are automatically associated with image schemas and abstract domains, presenting words related to image schemas and metaphoric expressions in a congruent font color should speed their categorization, while presenting them in an incongruent color is likely to hamper categorization speed (*Reverse Stroop Effect* (McLeod, 1991)). Study 8 therefore aims to complement the previous two experiments by studying the opposite direction of the implicit color association from color perception to semantic content. Focussing again on WARM-COLD as *source* domain, the hypothesis of this experiment is that literal and metaphorical meanings of WARM are categorized faster in *red* compared to *blue* font color and literal and metaphorical meanings of COLD are categorized faster in *blue* compared to *red* font color. German participants were recruited again since the stimulus material has to be carefully controlled for many variables and convenient lists of adjective norms do exist in German, e.g. (Hager & Hasselhorn, 1994), but not in Japanese.

7.3.2. Method

7.3.2.1. Participants and general procedure

The participant sample consisted of 31 (19 female, M_{age}= 20.9, SD_{age}= 2.86, range 19-29 years) undergraduates from Würzburg University, Germany, seeking course credit. All participants had normal or corrected-to-normal vision, and German was their mother tongue. Their favorite colors were blue (29.0%), green (22.6%), orange (12.9%), red (12.9%), black (9.7%), violet (3.2%), and other (3.2%). Identical to the procedure of the previous two studies, they received general instructions and then completed a demographic questionnaire, followed by a semantic categorization task on a personal computer.

7.3.2.2. Stimuli

Word material was collected to represent the WARM-COLD image schema and three of its metaphoric extensions (INTIMATE/EMOTIONAL/ACTIVE IS WARM - DISTANT/UNEMOTIONAL/PASSIVE IS COLD), see Tab. 7.4. For the two endpoints of each dimension, six synonyms were chosen. This was done to ensure the participants had to read and process the words in order to semantically categorize them, instead of just reading them aloud when a very limited word set is employed (Lorentz et al., 2016). In addition, following Lorentz et al., a larger word set can serve to test a larger portion of the semantic network and minimize repetition. Selected words were matched regarding length, number of syllables and word frequency, the latter again taken from the *Wortschatz* portal of Leipzig University[4]. Arousal and valence data from (Hager & Hasselhorn, 1994) are depicted for reference, but were not available for all stimuli. Another selection criterion was whether the words were associated with either *red* or *blue*. A pre-test with 10 participants determined *red/blue* associations and took the average assignments into account. Clearly, the participants matched synonyms of image schemas and their metaphorical extensions to either *red* or *blue* (last column of Tab. 7.4). The target words did not share the first letter with the colors (*red* HSB = 30,94,98; *blue* HSB = 240,66,88) in the response set and were presented in 20 point Arial font against gray background (HSB = 0,0,80).

7.3.2.3. Apparatus

The study was programmed and run using EPrime 2[5] on a 3.40 GHz computer with 4 GB RAM and a 60 Hz FUJITSU B24W-6 LED color monitor. Similar to the previous two studies, the monitor was placed about 50 cm in front of the participants who did use a common keyboard as input device. The response keys *H* and *J* were marked with white stickers.

7.3.2.4. Design and Procedure

The procedure was similar to Study 6 and Study 7. The participants completed a demographic questionnaire, followed by the main experiment. This time, their task was to make a bipolar semantic judgement on the presented words (WARM-COLD, INTIMATE-DISTANT, EMOTIONAL-UNEMOTIONAL, ACTIVE-PASSIVE). Speed and accuracy were emphasized and the subjects were instructed to only use the index finger of their dominant hand, see sec. A.8.2 for the instructions in German. The response key assignment was counterbalanced between participants and was re-instructed before each experimental block. The event sequence of each trial was identical to the previous two studies and is illustrated in Fig. 7.5. A black fixation point was shown at the center of the screen for 300 ms, followed by a blank screen for 500 ms and the stimulus for a maximum duration of 3000 ms or until a response was given by the participant. A prior practice block included 52 trials with nouns (*luck, enemy, love* etc.) that had to be categorized regarding their valence (*positive, negative*). Within this block response speed and accuracy were fed back to the participants. The following eight blocks of experimental trials contained 48 trials each (12 words per condition presented twice in *blue* and *red*, shown in random order without repetition), totaling in 384 trials. Each block (image schema, three metaphors) was repeated twice and the block sequence was randomized. A short break was offered after each block. The whole procedure took about 30 minutes.

[4]http://wortschatz.uni-leipzig.de/index.html, last accessed 18.09.2016
[5]www.pstnet.com

Table 7.4.: *Overview of the Characteristics of the Stimulus Material Used in Study 8 (Translated from German)*

Word	Condition	Temperature	Frequency	Length	Syllables	Arousal	Valence	Red/Blue
warm (warm)	image schema	warm	11	4	1	-0.7	1.24	1.00
heiß (hot)	image schema	warm	10	4	1	n/a	n/a	1.00
gewärmt (warmed)	image schema	warm	17	7	2	n/a	n/a	1.00
geheizt (heated)	image schema	warm	14	7	2	n/a	n/a	1.00
hitzig (fervid)	image schema	warm	15	6	2	n/a	n/a	1.02
glühend (burning)	image schema	warm	16	7	2	n/a	n/a	1.00
kalt (cold)	image schema	cold	10	4	1	-0.18	-1.06	2.00
kühl (cool)	image schema	cold	12	4	1	-0.21	-1.53	2.00
frisch (chilly)	image schema	cold	10	6	1	n/a	n/a	2.00
eisig (icy)	image schema	cold	15	5	2	n/a	n/a	2.00
frostig (frosty)	image schema	cold	15	7	2	n/a	n/a	1.97
schattig (shady)	image schema	cold	17	8	2	n/a	n/a	2.00
intim (intimate)	metaphor	warm	15	5	2	0.66	0.82	1.08
sexuell (sexual)	metaphor	warm	12	7	3	n/a	n/a	1.00
herzlich (sincere)	metaphor	warm	10	8	2	1.30	0.28	1.05
liebevoll (loving)	metaphor	warm	11	9	3	2.53	0.71	1.08
leidenschaftlich (passionate)	metaphor	warm	13	16	4	1.73	1.79	1.00
erotisch (erotic)	metaphor	warm	15	8	3	n/a	n/a	1.00
distanziert (distant)	metaphor	cold	13	11	3	-0.36	-0.40	1.92
fern (far)	metaphor	cold	11	4	1	-0.31	-0.40	2.00
reserviert (aloof)	metaphor	cold	12	10	3	-0.67	-1.14	1.90
verschlossen (reserved)	metaphor	cold	12	12	3	-0.56	-1.06	1.87
steril (sterile)	metaphor	cold	15	6	2	n/a	n/a	1.97
zugeknöpft (cagey)	metaphor	cold	16	10	3	-1.60	-1.36	1.92
aktiv (active)	metaphor	warm	9	5	2	1.46	1.12	1.08
tatkräftig (energetical)	metaphor	warm	13	10	3	2.20	2.64	1.08

Table 7.5.: *Overview of the Characteristics of the Stimulus Material Used in Study 8 (Translated from German) (Continued)*

Word	Condition	Temperature	Frequency	Length	Syllables	Arousal	Valence	Red/Blue
lebhaft (lively)	metaphor	warm	14	7	2	1.15	1.31	1.08
agierend (operating)	metaphor	warm	18	8	3	n/a	n/a	1.03
handelnd (acting)	metaphor	warm	18	8	2	n/a	n/a	1.08
energisch (vigorous)	metaphor	warm	13	9	3	0.53	1.08	1.00
passiv (passive)	metaphor	cold	13	6	2	-1.21	-1.08	1.97
tatenlos (inactive)	metaphor	cold	13	8	3	n/a	n/a	1.92
untätig (idle)	metaphor	cold	14	7	3	-2.07	-2.71	1.95
teilnahmslos (apathetic)	metaphor	cold	16	12	3	-1.93	-2.71	1.92
leblos (lifeless)	metaphor	cold	14	6	2	n/a	n/a	1.95
unbeteiligt (uninvolved)	metaphor	cold	15	11	4	n/a	n/a	1.97
emotional (emotional)	metaphor	warm	12	9	4	0.70	1.01	1.03
aufgewühlt (troubled)	metaphor	warm	15	10	3	n/a	n/a	1.10
gefülhlsmäßig (emotive)	metaphor	warm	16	12	4	n/a	n/a	1.08
erregt (excited)	metaphor	warm	12	6	2	n/a	n/a	1.00
empfindsam (sensitive)	metaphor	warm	17	10	3	1.12	-0.71	1.21
gefühlsbetont (sentimental)	metaphor	warm	19	13	4	0.99	-0.21	1.05
sachlich (objective)	metaphor	cold	12	8	2	-0.70	0.63	2.00
gefühllos (callously)	metaphor	cold	16	9	3	-2.33	-0.57	1.95
nüchtern (sober)	metaphor	cold	12	8	2	0.33	-0.50	1.95
emotionslos (unemotional)	metaphor	cold	15	11	4	n/a	n/a	2.00
rational (rational)	metaphor	cold	14	8	3	-0.33	0.64	1.97
vernünftig (reasonable)	metaphor	cold	22	18	3	n/a	n/a	n/a

Note. The word frequency class shows how much more frequent the German word “der” is compared to the target word in the form $2^{\text{frequency class}}$ according to Wortschatz Portal Universität Leipzig; Red = 1, blue = 2 (n=10).

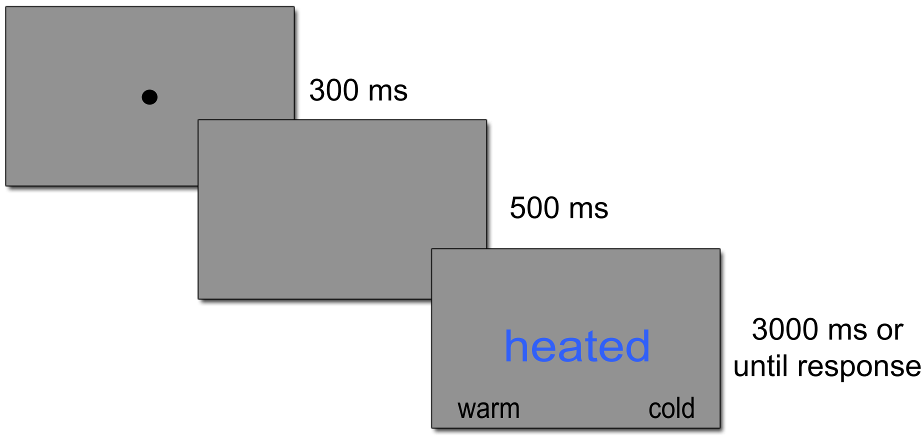

Figure 7.5.: Example of a trial display in Study 8.

7.3.3. Results

The same procedures as in Study 6 and Study 7 were applied to handle reaction times, including the deletion of inaccurate responses (5%), log-transformation and replacement of reaction times in trials which were 2.5 standard deviations below or above the grand mean of each condition (image schema/metaphor, congruent/incongruent) with these values. Responses faster than 300ms were omitted. The reaction times of the three metaphor conditions were averaged, as they did not differ, $F<1$.

A 2 (Condition: image schema, metaphor) x 2 (Congruency: congruent, incongruent) repeated-measures ANOVA was conducted on the (trimmed) log reaction time means, though millisecond means will be reported for ease of interpretation. There was a main effect for Condition, $F(1,30) = 98.18$, $p < .001$, $\eta_p^2 = .77$. Image schemas (M = 722.88, SD = 91.82) were categorized faster than metaphors (M = 833.26, SD = 122.89). The main effect of Congruency was also significant, $F(1,30) = 26.20$, $p < .001$, $\eta_p^2 = .47$. Incongruent trials (M = 789.12, SD = 101.86) were categorized slower than congruent trials (M = 767.02 SD = 106.80). Moreover, the Condition by Congruency interaction reached significance, $F(1,30) = 28.66$, $p < .001$, $\eta_p^2 = .49$. Millisecond means for this interaction are displayed in Fig. 7.6, suggesting that only for the image schema condition, semantic categorization was impaired if the word was shown in an incongruent font color.

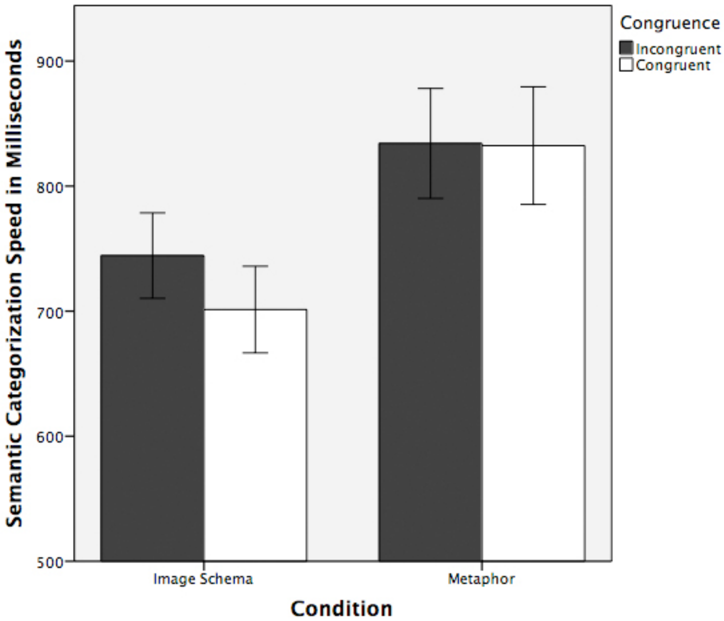

Figure 7.6.: Semantic categorization speed as a function of Condition and Congruency (means with 95% CI error bars).

Follow-up tests examined the effect of congruent vs. incongruent word-color mappings in each experimental condition separately, confirming that incongruent trials (M = 744.38, SD = 93.26) were categorized slower than congruent trials (M = 701.38, SD = 94.36) in the image schema condition, $t(30) = 6.10$, $p < .001$, $d = 0.46$, but not in the metaphor condition (incongruent: M = 833.86, SD = 119.63; congruent: M = 832.66, SD = 128.15), $t(34) = 0.46$, $p = .652$.

7.3.4. Discussion

The results of Study 8 indicate that response times were delayed when participants had to categorize the meaning of the WARM-COLD image schema dimension presented in incongruent font color, but not when they categorized the target domains of *temperature* metaphors. Overall, the results from Study 6, 7 and 8 suggest that there is no automatic activation of the three tested color substituted image-schematic metaphors - neither in the direction from semantic metaphorical content to color categorization speed nor from color to semantic categorization speed of metaphorical extensions of the WARM-COLD domain. Thus, the present results do not support implicit links between image schema-indicating colors and abstract domains, and contrasts with the findings that could be found under conscious processing in Study 5, where participants matched both in over 95% of the cases. This suggests that while color, especially when combined with temperature sensations, is reliably associated with metaphoric concepts that have WARM-COLD concepts in the source domain

(Study 5: 96%), this link is not established automatically. Therefore, color metaphors that have a color property in their source domain, e.g. ANGER IS RED, qualitatively differ from color substituted image-schematic color metaphors, e.g. ACTIVE IS RED.

Moreover, color metaphors like ANGER IS RED, SAD IS BLUE and BRIGHTNESS IS VALENCE do have counterparts in language and have an innate and/or experiential origin. Such color metaphors can be reliably identified in experiments involving subconscious processing (Meier et al., 2004, 2007; Fetterman et al., 2011; Okubo & Ishikawa, 2011; Lakens et al., 2012; Song et al., 2012; Fetterman et al., 2012; Young, Elliot, Feltman, & Ambady, 2013; Lakens et al., 2013; Thorstenson et al., 2015). Whether the association between color and abstract concepts in such metaphors is uni- or bidirectional or if there are some mediating factors is still a matter of scientific debate (Lakens, 2012), as empirical results exist in support of both views. Contrary, color substituted image-schematic metaphors do not have counterparts in language. Rather, the image schema functions as a source domain for abstract and color concepts alike (e.g., *warm red*; *warm thoughts*). Since color properties and abstract content rarely frequently correlate in the environment, it is not surprising that only those links are internalized and automatically applied which are reliably encountered, e.g., *blue* and SADNESS (Bubl et al., 2010; Barchard, Grob, & Roe, 2016).

Study 8 revealed an implicit effect of font color (*red* vs. *blue*) on semantic categorization speed of WARM-COLD synonyms, which was not obtained the other way around (Study 7). This suggests that color-image schema associations are automatically activated, fulfilling the precondition of intuitive use. There are at least two ways to account for such an asymmetric effect which seem to be compatible. First, as perception precedes conception, it is more likely for color perception to influence conceptions than vice versa (Rolls, 2000; Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006). Support for this notion can be found in that the reaction times of the semantic categorization task (699.83 ms on average for “hot” and “cool” in the image schema condition, Study 8) were overall much slower than the time required for color categorization (447.19 ms on average for the comparable words “cold” and “warm” in the image schema condition, Study 7). Image schemas were classified as *red* or *blue* much more quickly than they were semantically categorized. This supports the notion that perception is the easier achievement compared to semantic categorization (Fetterman et al., 2012). Second, the color-temperature association might simply be stronger in the direction of color to temperature than vice versa because colors are often used in our environment to indicate temperature (e.g. on water tabs), but temperature is rarely, if ever, used to indicate color (Ho, 2015).

Overall, semantic categorization was slower for metaphors than for image schemas (110 ms). Although the incongruent trials of the image schema condition have caused greater interference compared to congruent trials, the response times of incongruent image schema trials are still faster than for the metaphor condition (congruent, incongruent). This can be due to the fact that abstract metaphoric concepts require more time to be semantically categorized than physical properties.

One last aspect that needs to be briefly mentioned are possible cultural differences in implicit associations between color, temperature and temperature metaphors. In Study 6, Japanese subjects categorized font color of temperature-related words. It could be found that reaction times for the (incongruent) metaphor trials were slower than for the (incongruent) image schema and (incongruent) color trials. Although a regression analysis did not yield a considerable impact of differing word length and word frequency between the conditions on the results, it was nevertheless a confounding variable. Study 7 therefore replicated the experiment with minor changes, including the introduction of neutral trials and congruent conditions as a control, as well as using German participants. Although the comparison between congruent and incongruent trials within each condition (color, image schema, metaphor) was in focus of the analysis, the incongruent trials of each condition still showed the same pattern as with the Japanese subjects: incongruent color trials were slowest, followed by incongruent image schemas and then incongruent metaphoric expressions. Nonetheless, the studies conducted within this work can not fully determine to which extent this difference in reaction times is due to different associative strengths or due to differences in lexical characteristics of the stimulus material. Study 8 took a closer look on the opposite side of the relationship, whether color influences semantic categorization speed. This study was not replicated with Japanese participants, mainly because there is currently no database available for Japanese with detailed information about word frequency, arousal and valence, and the development of such a database is out of the scope of this work. However, as Germans and Japanese are exposed to similar temperature stimuli in their daily lives, similar results can be expected for both cultural backgrounds (Ho, 2015).

Finally, the obtained results could also be attributed to the Stroop paradigm and the limitation to temperature associations. The Stroop test was chosen because both dimensions of interest are integrated in a single stimulus (e.g., colored text), which represents a common use case in HCI. Moreover, the Stroop test is a widely used tool to access interference effects between colors and semantic content and has been replicated in a couple of hundred studies (McLeod, 1991). Needless to say, other implicit measures of associative strength exist, like the *Implicit Association Test* or priming tasks. Although not including temperature metaphors, other authors already investigated color-image schema associations using these experimental paradigms, gathering similar results. For example, Ho and colleagues used the *Implicit Association Test* and found that *red* and *blue* are automatically associated with *warm* and *cold*, respectively (Ho, 2015). Their second experiment consisted of a priming task, in which the semantic condition (priming color-related words with color patches and vice versa) showed a null-result because of a very long 2000ms stimulus onset asynchrony. However, in the other condition, physical temperature cues were used instead of thermal words. The authors only found a priming effect when color patches were used as primes and physical temperature as target, but not the other way around, supporting the results found in this work that colors do automatically activate the WARM-COLD image schema, but not the other way around. Moreover, as different measures to assess the strength of automatic associations between different concepts often yield similar results (Snowden, Wichter, & Gray, 2008), it seems not very likely that the results of the Stroop paradigm are only an artifact of the experimental method.

Another possibility why no implicit associations between colors and temperature metaphors were found could be that temperature metaphors are not automatically triggered, not even by physical temperature sensations. However, a variety of empirical results already demonstrated that the perception of temperature cues automatically activates the concept of *psychological warmth* (tested by INTIMATE-DISTANT) and vice versa (L. Williams & Bargh, 2008; Vess, 2012; Bargh & Shalev, 2012; Sun, 2012; Fay & Maner, 2015). Yet, a separate investigation of this metaphor did not show any differences compared to the other two metaphors (ACTIVE-PASSIVE, EMOTIONAL-UNEMOTIONAL). It is therefore concluded that there might be a difference between the automatic activation of temperature metaphors by physical temperature cues compared to thermal words. This question remains a topic for future research in this area.

7.4. Chapter 7 Recap

The three studies in Chapter 7 investigated whether color associations with the WARM-COLD image schema and related abstract domains are automatically activated and whether this activation occurs uni- or bidirectionally. No interference or facilitation effect of image-schematic and metaphoric semantic content on font color categorization speed compared to control conditions could be found (*Stroop Effect*, Studies 6 and 7). However, color categorization performance was overall slowest for incongruent color words, faster for image schemas and fastest for metaphors. While this result can be explained through the semantic gradient between the conditions, an alternative explanation is the difference in word frequency and length. A regression analysis quantified the impact of the latter, yielding that the semantic gradient played a more important role than frequency and word length. However, this confounding variable leads to an overestimation of the influence of semantic content on the speed of color categorization.

When examining the other direction of influence of the color-temperature association (Study 8), an implicit effect of font color on semantic categorization speed of image-schematic content could be found, but not for metaphoric meanings of temperature-related words. When font color mismatched presented image schema synonyms, reaction times were slowed down compared to when font color matched semantic content. This interference effect is likely to occur at the post-perceptual level of response selection (Melara, 1989). In cases when font color and semantic content mismatch, color invokes a different response tendency than the presented words, which results in interference. Attention is not involuntarily shifted to stimuli which are not in line with our current set of goals (except in situations which are potentially harmful for the organism) (Folk, 2016). Therefore, Folk suggests, as the to-be-categorized word carried color information which was diagnostic of relevance (i.e., related to temperature), the attention system might have used this color information to enhance all word stimuli (described as voluntary feature-based selection) and subsequently guide attention to those stimulus features (described as voluntary space-based selection).

Overall, implicit color-temperature associations were found to be asymmetric: font color influences semantic categorization speed, but semantic content does not influence the speed of color categorization. Two possible explanations were discussed that can account for this unidirectionality: perceptual color judgements are faster than conceptual judgements of image-schematic or metaphoric content, rendering an influence of color perception on conception more likely than the other way around. In addition, color is used in our environments to indicate temperature, which might have strengthened this direction of the association. Linguistic metaphors indicate an association between color and the WARM-COLD image schema, but the predicted direction (semantic content influences color categorization) is opposite to what was found in the empirical studies described in this chapter. As previously discussed by Casasanto, linguistic metaphors are suitable to generate hypotheses about underlying cognitive relations, but cannot replace empirical tests to identify their cognitive reality (Casasanto, 2009b).

In Studies 7 and 8 no implicit effect of temperature metaphors on color categorization speed or color perception on semantic categorization speed of temperature metaphors could be found. It is concluded that implicit associations between colors and abstract concepts are only established when both are frequently co-experienced and subsequently surfaced in language. On the contrary, when colors experientially covary with image schemas that ground abstract concepts, colors are likely matched to these concepts under conscious processing (Study 5), but have no or only a practically irrelevant impact in situations involving subconscious processing. Image schemas can be characterized as adjectives describing physical experiences in their literal meaning which can be used to describe more abstract concepts in a metaphorical sense. It is therefore not surprising that the image schema-color association is much stronger than the association between color and metaphorical derivate. In addition, color is highly context dependent and conveys different information in different contexts (Elliot & Maier, 2012a). Thus, the same perceptual referent is co-opted for multiple purposes (Fetterman et al., 2011) and image schemas, weakening specific relationships of color substituted image-schematic metaphors.

What can be learned from the studies of the past two chapters for the application of color in HCI? Color associations with the WARM-COLD image schema are established automatically in a way that the perception of image schema-related colors facilitates the processing of image-schematic semantic content. This implicit link therefore meets the precondition of intuitive use as subconscious application of prior knowledge. If other image-schematic color associations operate in a similar way, designers can utilize color as predicted by the regression models developed in Study 1 to support the semantic content they want to convey, even when users do not consciously think about the color design. When designing applications in which image schematic domains are indicated by color, it should be avoided to present incongruent visual and haptic information, like presenting high temperatures and the color blue together, especially in interfaces where users have to respond to temperature cues (Ho, 2015). Incongruent presentations will result in prolonged processing time, which can be problematic in for example time- and safety-critical situations. Compared to that, incongruent presentations of color and abstract concepts do not seem to hamper performance implicitly, but may impair the holistic interaction experience when users consciously think about this mapping.

8. Insights for User Interface Design

Color guides for designers are replete with recommendations about which colors to choose when designing user interfaces. Often, such guidelines claim that by choosing the 'right' colors, the design will affect the user's energy level, creativity, productivity, mood and focus. For example, *blue lowers heart rate*, *green inspires innovation* and *orange generates enthusiasm*[1]. Such naive beliefs about inherent effects of and associations with colors are widely popular and are continuously promoted by consulting agencies and the media, but have received little to no scientific investigation and support, rendering such statements, at best, speculative. According to recent scientific advancements in color psychology, colors do not have an inherent meaning, but exhibit their influence depending on *context* (Elliot & Maier, 2012a). However, although 'color' is extensively studied with regard to physics, physiology, linguistics and categorization, appearance phenomena, deficiency and reproduction, in general, very little theoretical work has been done to account for semantic color associations. To amend this situation, Conceptual Metaphor Theory of Color (CMToC) was developed in this work, which predicts and explains semantic color associations that are shared by a large group of people in a large variety of contexts.

8.1. Predicting and Explaining Semantic Color Associations: Conceptual Metaphor Theory of Color

To estimate the contribution of CMToC, current theoretical considerations in color psychology were reviewed, compared in terms of their predictive and explanatory power regarding semantic color associations, and analyzed for weaknesses. Three theoretical approaches have been selected and introduced in sec. 2.2 based on the fact that they take a - more or less - holistic view on color psychology. Color-in-context theory offers widely applicable premises, like for example that the influence of color on psychological functioning is typically automatic and that color meanings are grounded in biologically based response tendencies to particular colors in particular situations and learned associations (Elliot et al., 2007; Elliot & Maier, 2012a, 2012b). However,

[1] https://www.quill.com/blog/workplace-culture/productive-palette-how-color-can-influence-workplace-behavior.html?cm_mmc=NEW_Infographic_GHG, last accessed 09.10.2016

in specific contexts, these principles are "only vaguely suggestive of precise hypotheses" (Elliot, 2015, p.2). The second discussed approach was Affective Meanings Systems (Osgood, 1969; Adams & Osgood, 1973), a parsimonious explanation for a variety of color associations. In general, Affective Meanings Systems predicts that an association is likely to occur when colors and concepts overlap in factors of affective meaning. Nonetheless, it is not explained why colors and other concepts acquire these semantic features in the beginning and the scope is limited to the affective component of meaning (Osgood et al., 1975). As a third work, the Ecological Valence Theory of Human Color Preferences was discussed (Palmer & Schloss, 2009). Palmer and Schloss stress the importance of individual color preferences for decision-making, in that people prefer to surround themselves with objects and environments in their favorite colors, e.g., they paint the walls in their living rooms in colors they like, or buy objects in their favorite colors. As the focus of this theory is very specific, it does not clearly connect to a broader understanding of the variety of color phenomena, like the influence of color on face perception, in achievement contexts, or on the perceived attractiveness of opposite-sex partners in romantic contexts, to name a few. It is therefore not surprising that Elliot, in a recent review on theoretical and empirical work in color psychology, expressed the need for "mid-level theoretical frameworks that comprehensively, yet precisely explain and predict links between color and psychological functioning in specific contexts" (Elliot, 2015, p.2). In this work, CMToC was developed to take position as such a mid-level theoretical framework.

CMToC is based on the idea from Conceptual Metaphor Theory (CMT) that the study of metaphorical language provides valuable insights into our mental models involving color. CMT states that we do use metaphors to explain one thing in terms of another that is literally unrelated but figuratively similar (Lakoff & Johnson, 1980). Statements like *grasp the idea* are figures of speech in which we use a phrase to suggest likeness to another. Metaphors are so ubiquitous in discourse that Lakoff and Johnson argued that metaphoric language does not only represent the way we talk, but also the way we think. Most fundamentally, abstract concepts are understood in terms of more physical experiences. In the example mentioned before, *understanding* borrows from the physical act of grasping something. In this work, this notion was applied to color psychology.

Colors appear in metaphorical speech mainly in two forms. First, in color metaphors, color in the *source domain* is mapped onto another, more abstract, *target domain.* Linguistic example expressions for such a color metaphor are *seeing red* or *feeling blue.* Such color metaphors can reveal how conceptual knowledge is represented involving color. Color metaphors have received considerable interest in the scientific community so far (sec. 4.2). They are likely motivated by co-experiencing color and certain states or emotions, like for example depression and impaired vision on the blue-yellow axis as well as impaired contrast sensitivity (Bubl et al., 2010). However, studies almost exclusively focussed on the colors black, white and red (Meier, 2016). In addition, the origin of color metaphors has to be carefully identified. Linguistic metaphors like *green with envy* or *the red thread* also structure symbolic concepts by means of color, but do not originate from sensorimotor experience. Rather, they are learned in the course of language acquisition in a specific cultural setting. Therefore, they are highly dependent on culture and are most likely not automatically associated, thus offering no potential for the design for intuitive use.

In the second type of metaphor colors are structured as *target domain* by other physical sensations (image schemas) in the s*ource domain*. For example, in expressions like *heavy blue* or *warm red*, physical weight and temperature concepts are mapped onto color properties, describing dark and reddish/yellowish tones, respectively. Such expressions point to possible associations between colors and physical properties and have received substantial support in linguistic as well as in behavioral studies (sec. 4.3). Image schema-color metaphors are supposed to originate from innate and sensorimotor levels of prior knowledge and thus offer potential for the design for intuitive use regarding physical information.

The third type of conceptual metaphor that was identified within the scope of CMToC did not receive a lot of theoretical and empirical attention so far: color substituted image-schematic metaphors. These metaphors are created from image-schematic metaphors (*giving someone the cold shoulder* - DISTANT IS COLD) that do have an image schema in the *source domain* that is often co-experienced with specific color attributes and can therefore be substituted by them (sec. 4.4.1). Image schemas are non-linguistic, pre-conceptual patterns of bodily interactions with the external world and scaffold more abstract concepts through the process of metaphorical mapping. By unveiling the relationships between colors and image schemas, it is possible to predict associations between colors and more abstract concepts linked by the underlying image schema. Since associations between colors and image schemas are often innate or universally acquired, they build on the lower stages of the prior knowledge continuum (Hurtienne & Blessing, 2007) and are expected to be rather stable between different cultures. Thus, color substituted image-schematic metaphors offer potential for the design for intuitive use regarding abstract information.

To summarize, unlike the other theoretical approaches in color psychology, CMToC makes concrete predictions for semantic color associations for a variety of physical as well as abstract domains that can inform design decisions in both hard- and software user interface design. Within its scope, origins of color associations have been categorized in four levels of prior knowledge: innate, sensorimotor, cultural and expert knowledge. Examples for color associations anchored at each level were provided along with hypotheses about the cultural stability. Focussing on image schema-color metaphors for conveying physical information and color substituted image-schematic metaphors for conveying abstract information, the predictions of CMToC were tested in eight empirical studies. Following the three research questions, the outcomes and conclusions are summarized in the following sections.

8.2. Conveying Physical Information with Color

The first research question addressed the use of color for conveying physical information like *weight* in user interfaces. Four online surveys were conducted that investigate the relationship between 16 image schemas and the color attributes hue, saturation and brightness in a total of 295 German and Japanese subjects. Study 1 and 2 provided estimations on the relative impact of color attributes on the strength of the association with an image schema in form of regression models. Moreover, it was demonstrated that the found associations did not substantially vary between the two different cultures. The results were explained in terms of color-in-context theory, Affective Meanings Systems, Ecological Valence Theory of Human Color Preferences and CMToC, the latter having the highest accuracy in predicting and explaining the observed phenomena (79.2%). Studies 3 and 4 aimed at replicating the outcome of the previous surveys and extended the experimental setting from single color patches to two-color comparisons. Color pairs were selected based on the regression values of Study 1, and except for BIG-SMALL, the color pairs were assigned with 'almost perfect' agreement to the image-schematic poles. Again, the results between Japanese and German populations were very similar. It was concluded that CMToC provides the most comprehensive explanation for image schema-color associations and generates testable and falsifiable predictions regarding which colors go best with different physical properties across different cultures.

These findings can be systematically used when conveying physical information in user interfaces. Color is a very powerful means that conveys information about material properties prior to other senses like touch, smell or taste. These learned associations and long held expectations influence, alter and even trump the cognitive processing of other sensory experience in a top-down fashion (Garber, Hyatt, & Nafees, 2015). For example, by looking at an object, judgements about weight are consistently and reliably based on vision (C. Taylor, 1930). If an object looks like metal, it probably is - all along with the associated expectancies of heaviness, like an initial calibration of grip and load forces (Buckingham, Cant, & Goodale, 2009). When lifting the object, the motor system quickly adapts to potential misconceptions based on vision, but the overall expectation about the relationship between color, object density and weight remains persistent. This top-down control by visual cues explains why phenomena like the brightness-weight illusion remain unchanged and moderate perception even after being exposed to several inconsistent instances (P. Walker et al., 2010). However, the impact of color is different between the senses. Some physical stimuli trigger innate responses, like reactions to basic tastes, offering less scope for color influences, while other associations are learned and are thus prone to color effects (Shankar et al., 2010). In any case, color can help to disambiguate material or object identity and, when appropriately used, contribute to a homogenous user experience regarding physical properties. Therefore, by knowing which color attributes influence the perception of other physical sensations, designers can play with the users' expectations of certain functionalities or properties, like creating surprise reactions through amusing visual-tactual incongruities (Ludden et al., 2009, 2012; Ludden & Kudrowitz, 2012).

8.3. Conveying Abstract Information with Color

The second research question addressed the use of color for conveying abstract information like *importance* in user interfaces. In one experimental study, a total of 75 German and Japanese participants validated color-to-abstract mappings in form of color population stereotypes like *important is dark*. The participants rated translation-equivalent words like *happy*, *significant* or *guilty* to visual and haptically instantiated image-schematic objects by simply pointing to that alternative which seemed to go best with the word in question. Following CMToC, predictions were made which objects will likely be matched to the presented adjectives based on color substituted image-schematic metaphors. Participants frequently (88%) assigned congruent objects to corresponding metaphorical extensions of the image schemas BIG-SMALL, HEAVY-LIGHT, STRONG-WEAK and WARM-COLD. Words opposed in meaning (e.g., *happy* and *sad*) typically displayed reciprocal profiles against the image-schematic polarities. Both cultures agreed significantly on what aspects of meaning were related to what dimensions of visual and haptic experience. These findings were interpreted as successful extension of CMToC to relationships between color and more abstract concepts that are not expressed in language. Moreover, when similar image-schematic metaphors could be found in Japanese and German, it is likely that subjects cross-culturally link specific color properties of objects to structurally corresponding symbolic concepts, moderated by the underlying image schema.

In total, Study 5 tested color associations with 20 abstract concepts of four image-schematic domains. The results suggest that following color substituted image-schematic metaphors may inform the design in such a way that it affects the experience of using physical artifacts (Löffler, 2014a, 2014b). When colors are aligned with image schemas they naturally indicate, or even substitute them, participants match colored objects more reliable to abstract concepts (Löffler et al., 2016), need less time, and indicate higher preference ratings. It can be concluded that designers therefore need to explicitly design for color, because image schema-incongruent colors can hamper the effectiveness of metaphorical mappings. Comparisons between different colors that similarly represent an image schema suggest that it is not specific color variants but the dimensional relationship that drives the effectiveness of the mapping, thus circumventing to argue about specific colors with high-level symbolic meanings. Moreover, own studies have shown that color-for-haptic substitutions work equally well in a digital environment (Löffler et al., 2015). Designs that draw on such experience-based meaning of color may come in conflict with more conventional but sometimes arbitrary ways of expressing physical and abstract information with color. Therefore, following CMToC in design can be understood as striving towards more familiar experiences of information technology on a deeper level of prior knowledge.

8.4. Cross-Cultural Validity of Semantic Color Associations

All three types of conceptual metaphors identified in this work that are fruitful for the prediction of semantic color associations are located at the lower levels of prior knowledge (innate, sensorimotor). Therefore, the likelihood that they would generalize across cultures is high. This is true for associations between color and other physical sensations, as well as their metaphorical extensions which are hypothesized to derive from correlations in basic sensorimotor experience. As many physiological processes that are involved in the formation of image schemas are the same across cultures and even across species, the notion that color associations with image schemas are less variable across cultures seems rather intuitive. But how about more abstract and intangible concepts? In linguistics, cross-cultural variation of conceptual metaphors is a well-investigated topic. However, cross-cultural studies going beyond the analysis of discourse are rather scarce.

In this work, three different types of empirical studies were conducted with subjects of different cultural backgrounds. In the surveys (Studies 1-4) as well as the population stereotype experiment (Study 5), a total of 370 subjects with either German or Japanese as their mother tongue matched colors to image schemas or metaphorical extensions of image schemas. For image schema-color associations, predictions of associative ratings of Japanese based on data of German participants strongly correlated with the empirical data across all image schemas. When colors substituted an image schema (Study 5, *color* condition), Japanese preferred them slightly more than the German participants and were slower in subjecting their responses. The slower response times were a systematic pattern which could be observed across all conditions. Speculating, this can be attributed to a slow decision making processes in Japanese society in general as a consequence of strict hierarchy, high uncertainty and risk avoidance as well as decision making by consensus (Hofstede, 2001). However, participants of both cultures matched the colored objects to the abstract concepts in a very similar way, which supports the assumption that color associations on the lower levels of prior knowledge, i.e., innate and sensorimotor knowledge, are valid for a large range of people with different cultural backgrounds.

8.5. Colors and Design for Intuitive Use

Intuitive use is defined as the extent to which a technology can be used by subconsciously applying prior knowledge (Naumann et al., 2007). The application of knowledge can become subconscious "due to frequent exposure and reaction to stimuli in the environment: the more frequent the encoding and retrieval was in the past, the more likely it is that memorized knowledge is applied automatically and subconsciously" (Langdon, Clarkson, & Robinson, 2008, p. 108). The last research question addressed whether image schema-color metaphors and color substituted image-schematic metaphors are processed automatically as a precondition for intuitive use. Studies 6-8 employed a Stroop-like paradigm to study the directionality and automaticity of both metaphoric mappings. While participants had no time limit to make their responses in Studies 1-5, the last three experiments imposed a time restriction. In a speeded forced-choice task, subjects had to categorize the font color of linguistic instantiations of the WARM-COLD image schema and related metaphorical extensions (Studies 6 and 7) or semantically categorize the word meaning of colored words (Study 8). It could be shown that perceiving color automatically influenced the identification speed of related physical concepts, but not of metaphorical content. In the other direction (influence of semantic content on font color categorization speed), no effect of image-schematic or metaphoric content could be observed compared to control words. However, the speed of color categorization showed a typical semantic gradient effect (Klein, 1964), meaning that the stronger the association between word content and color, the slower the reaction times in the incongruent condition. Color categorization performance was slowest for color words, faster for image schemas and fastest for metaphors. This effect is partly superimposed by differences in word frequency and length between the three word types. Overall, color-temperature associations were found to be automatically activated in an asymmetric way: a mismatching font color slows down categorization speed of image-schematic semantic content, but font color categorization is unaffected by image-schematic semantic content. No effect of metaphoric temperature metaphors on color categorization or the other way around could be observed.

Two main conclusions can be drawn from these findings. First of all, implicit and explicit color associations can differ. In Study 5, participants reliably matched colored objects to corresponding metaphoric *warm-cold* concepts (86% accuracy in congruent color trials across German and Japanese subjects) as predicted by color substituted image-schematic metaphors. In Study 8, no implicit and automatic effect of color on semantic categorization of *warm-cold* metaphors could be found. Compared to this, evidence for implicit and explicit image schema-color associations as indicated by image schema-color metaphors was found. Regarding color metaphors that are based at the sensorimotor level of prior knowledge, like ANGER IS SEEING RED, evidence exists for both for explicit associations (for example, see 'symbolic' color associations common in media) as well as implicit and automatically activated associations (Fetterman et al., 2011, 2012; Fetterman, Liu, & Robinson, 2015). Contrary, color associations based on the cultural level of prior knowledge, like *red thread* used in a metaphoric sense for structure, are explicitly used in discourse. If they also implicitly and automatically affect conceptual choices is a topic for further research, as no studies have investigated such effects so far, at least to the extent of the author's knowledge.

The second conclusion drawn from the series of studies concerns the question of symmetry of color associations. As investigated on the example of the WARM-COLD image schema in Studies 6-8, the association between image schema and color is asymmetric. While color does influence semantic categorization of image-schematic content, the opposite does not hold true: no automatic effect of semantic content on color categorization could be observed. Whether this basic asymmetry is unique to the WARM-COLD image schema or can also be found in other image schema-color associations and other experimental paradigms is an area for future research. Regarding conceptual color metaphors on sensorimotor basis, the results have been mixed. Some associations like ANGER IS SEEING RED are asymmetric in a sense that perceptual redness influences semantic categorization of anger-related concepts, but not vice versa (Fetterman et al., 2011, 2012). Other associations have been reported to be bi-directional, like for example BRIGHT IS GOOD (Meier et al., 2004, 2007; Okubo & Ishikawa, 2011; Lakens et al., 2013). The factors that influence the directionality of such color associations remain a question for further research.

Summarizing, according to the findings of this work and the reviewed literature, when colors are part of metaphoric expressions in the target domain (image schema-color metaphor) or source domain (color metaphor), they are explicitly and implicitly associated with the *source* or *target domain*, respectively, because they originate from the lower levels of prior knowledge and are frequently co-experienced. Color substituted image-schematic metaphors, on the contrary, are not expressed in linguistic metaphors, but are still explicitly recognized. Whether or not the experiential correspondence on a sensorimotor level is essential for the formation of implicit and automatically activated color associations, or if color associations can also be established at a subconscious level and automatically retrieved when acquired solely through language or other cultural practices is still an open question that requires further research.

Overall, CMToC is a pragmatic theory that allows for designing solutions that are better rooted in our cognitive capacities and are thus more implicitly meaningful. Building on color associations that people are naturally familiar with (Raskin, 1994) supports the subconscious application of prior knowledge, and, as a result, intuitive use (Hurtienne & Blessing, 2007). In addition, if users can rely on prior knowledge about similar use of color in other tools, it will decrease the time and errors in interacting with new devices (Hurtienne & Langdon, 2009). By following empirically validated conceptual metaphors involving color, designers can reduce the ambiguity of color through established mental models and deliver coherent sensory messages. Chapters 4-6 reported relevant literature, surveys and experiments that quantify the relative importance of color attributes for literal and metaphorical meanings of sensory descriptors. Taken together, those color associations that automatically established can inform the design of intuitive-to-use interfaces by means of color.

8.6. Future Directions

This work provided a comprehensive overview of how the color properties hue, saturation and brightness relate to 16 image schemas, commented on where the association originates from and derived predictions about cultural stability from this. The predictions were tested in a Western and Eastern culture, German and Japanese, respectively. In addition, color instantiations of image schemas in selected primary metaphors of the physical domains of WARM-COLD, STRONG-WEAK, BIG-SMALL and HEAVY-LIGHT were cross-culturally examined. The results are encouraging in a way that color associations like those identified within the framework of CMToC do allow concrete predictions in a variety of physical as well as abstract domains, precisely because they seem be consistently acquired and less variant across cultures. It remains a topic for future research to extend and validate this experience-based approach in other cultures, especially in those who live in significantly different (natural) environments and are thus exposed to different statistical regularities of color, like for example, a tribe living in jungle or the Inuit. This work has taken a first step to provide a theoretical and empirical basis for cross-cultural semantic color associations, but work in color psychology is still at an early stage of development (Elliot, 2015). Several opportunities for further work are discussed in the following.

8.6.1. Empirical Work

One big issue that arises when conducting psychological color experiments is the question of *internal vs. external validity*. Generalizability can be increased when constricted environments like in this work are employed that study specific effects in isolation and minimize context effects. Indeed, many factors have to be considered when presenting color stimuli to participants: measures of color control, viewing distance, angle, amount and type of ambient light, presence of other colors in the immediate background - to name some of them. The practical relevance of these factors in UI design is an understudied area. For example, as the results of Studies 1-4 have shown, the BIG-SMALL image schema and related color associations are highly depending on foreground-background color contrast. It is therefore necessary to have strictly controlled experimental conditions and researchers reporting all experimental details in their publications in order to sharpen our understanding of the relative importance of environmental factors and to decrease interpretational ambiguity of the results (Elliot, 2015). On the other hand, color perception is a complex phenomenon and we are facing a multitude of colors in user interfaces. Thus, studying color only in strictly controlled environments is far from the complex interplay between color and surrounding factors in real-world scenarios. To balance this trade-off between internal and external validity, more 'ecologically valid' experiments are needed (Lewin, 1951; Barker, 1978; Brunswik, 1956) that investigate color meaning effects embedded in user interfaces to estimate the transferability of laboratory results to applied contexts.

The questionnaire studies on which the empirical part of this thesis is based were conducted in real-world settings. Subjects participated online and accessed the to-be-rated color patches on their private computer screens. Although the background color was kept constant and they were asked to maximize monitor brightness, display settings were not controlled for, as well as other parameters like viewing distance, viewing angle, or the amount and type of ambient light. Still, the split half reliability between the obtained ratings of the participants as measure for internal consistency was above .90 in all cases. Findings of other authors are also in support of the relative stability of semantic color associations across different viewing conditions. For example, Taft reported that almost identical semantic ratings can be obtained when evaluating a colored chip compared to a colored object (Taft, 1997). Given that one aim of this thesis is to guide color design in HCI, studies with higher external validity as those conducted within this work might have higher practical value for the field.

The two types of empirical studies conducted in this work (questionnaires, experiments) show a trade-off between the amount of work involved and the effectiveness of the research. In large scale studies many color associations can be investigated (Studies 1-4), whereas for clarifying specific aspects of semantic color associations, only a few can be taken into account (Studies 5-8). If the goal is to quickly build a knowledge base for designers, the former is recommended. If boundary conditions like automaticity and directionality of color associations are in focus of the research like in the last three studies, more fine-grained experiments are required that only focus on a few or even a single semantic color association.

Now that some premises of CMToC have been put to empirical test, the door is open for more applied studies. These can include the study of more diverse cultures, especially those living in different natural environments, as well as larger sample size studies (Elliot, 2015). The results in this work should be treated as preliminary findings awaiting replication in subsequent work in order to develop a robust scientific literature (see Elliot 2015 for a deeper discussion of the issue of small sample sizes). Besides a replication in other cultures and with larger and more diverse sample sizes than university students, other image-schema color associations (e.g. with categorial image schemas line IN-OUT, BLOCKAGE etc.) and color substituted metaphors should be included, as well as other methods. In addition to the above mentioned research topics, it might be worthwhile to have a look into how dynamically changing colors in our environment and user interfaces might affect the attribution of meaning, e.g. a night mode that changes the screen's luminosity depending on the time of the day. So far, research is limited to studying static colors and, also much less frequently, color transitions.

8.6.2. Theorizing Work

Much on color meaning in HCI has mainly focussed on hue. As the theoretical work and the empirical studies within this work have shown, it is worth *turning the* *attention to saturation and brightness* as well (Elliot & Maier, 2014; Elliot, 2015). The work conducted in this thesis therefore represents a first step taken into this direction, as the three properties hue, saturation and brightness have been equally investigated. Another factor that needs to be considered towards a more complete understanding of color effects in HCI is the identification of *mediating and moderating* factors, like contextual cues, situational aspects, as well as task types (Elliot, 2015; Meier, 2016). By moving from asking questions like "does color X influence Y?" towards asking questions like "when does color X influence Y?", boundary conditions of semantic color associations can be identified that broaden applicability. Boundary conditions studied within the scope of this work are cultural background (Studies 1-5), cross-modal effects (Study 5) and direction of the association (Studies 6-8). It was found that the cultural background of German and Japanese subjects had little influence on color associations predicted by image schema-color metaphors and color substituted image-schematic metaphors. With regard to modality, subjects matched colored objects to metaphorically linked abstract concepts faster when they were not required to touch them, but also slightly less reliable. In terms of task type, it could be found that when subjects had to classify image-schematic semantic content, their response times were influenced by the font color in which the stimuli were written. If their task was to classify font color, image-schematic and metaphoric semantic content did not have an influence on the speed of responses. It was concluded that color associations of the temperature domain operate unidirectional.

Another blind spot in research is the study of *interpersonal and intrapersonal variables*. Although semantic color associations based on innate and sensorimotor levels of prior knowledge are assumed to be acquired universally, differences in color perception might alter the effects of color in HCI for the individual user. For example, Fetterman et al. showed recently that higher preferences for red covary with more hostile social decision-making (Fetterman et al., 2015), boosting metaphor-consistent behavioral consequences for specific users. Another adjacent research issue are visual impairments. Future work could examine if the link between color perception and metaphoric concepts is weakened or non-existent for people with impaired color vision. CMToC predicts different semantic color associations if the experiential basis is altered.

8.7. Conclusion and Outlook

8.7.1. Application in User Interface Design

Admitting, the line between appropriate and inappropriate use of colors is very thin, but if used properly, colors can be a powerful tool to align UIs with the users' mental models. When designing UIs, various choices regarding layout and color have to be made. Besides their own intuition, designers can draw on diverse design guidelines and common beliefs which colors do best represent certain concepts. For example, best practices exist how to depict frequently used functions or signs using color, e.g. safety-related signals like warnings (Wogalter, Conzola, & Smith-Jackson, 2002). However, as color is ubiquitous and very easy to implement and change, it is oftentimes also arbitrarily applied. According to Humphrey, such indiscriminate use of color on man-made objects dulls our biological predisposition to interpret it as a meaningful signal (Humphrey, 1976). Therefore, in order to unveil experience-based couplings between color and meaning, CMToC was developed within this work. Following this, origins of color associations were described on the four different levels of prior knowledge from innate to sensorimotor, to cultural up to expertise (Hurtienne & Langdon, 2009). Many color associations identified through analyzing discourse in different languages primarily rest on the lower levels of prior knowledge. If UIs build on semantic color associations that are grounded at the innate and sensorimotor level, a large range of people of different cultural backgrounds will be able to make sense of them, since they are so basic and fundamental. Regarding UI design, this means that less work has to be put into localizing color in design variants for different cultures. In addition, according to Hurtienne and Langdon, sensorimotor knowledge is so frequently encoded and retrieved throughout lifetime that it should be less sensitive against individual differences in cognitive abilities and technology familiarity. Basing the design on innate or empirically substantiated and widely established sensorimotor color associations instead of relying on color associations on the cultural and expert level of prior knowledge can also have another advantage. By designing tools, objects and interfaces in a way consistent with nature's color-coded messages constrained by our physical bodies, we can shape color experiences with man-made artifacts and create consistent correspondences both within and across cultures. This would strengthen the structure metaphors impose on our lives, and serve as an experiential basis for these metaphors in the next generation of people (Lakoff, 1993).

This work functions as an encyclopedia about semantic color associations on innate and sensorimotor stages of prior knowledge and centralizes knowledge that is currently scattered throughout the literature. As only some of the exemplary derived conceptual metaphors involving color have received empirical support so far, they can inspire designers and researchers to generate new hypotheses on color associations with physical and concepts. Going back to the example of online banking mentioned in the introduction of this work, a designer can now determine which colors go best to convey the concept of *security*, which is important when handling sensitive data and money via the internet. For this, linguistic metaphors involving *security*, and, if possible, color, need to be found first. In conventional speech, no metaphors directly link color to *security*. Therefore, a color substituted image-schematic metaphor needs to be constructed as described in sec. 4.4.1. Color substituted image-schematic metaphor are a combination of image schema-color metaphors and image-schematic metaphors. The first are documented in Tab. 4.1 and many of the latter in ISCAT, a database containing image schema instances in language and user interfaces [2]. Image-schematic

[2]http://zope.psyergo.uni-wuerzburg.de/iscat, last accessed 15.03.2017

metaphors related to the concept of security are for example SECRET IS HEAVY (*heavy secrets*) and PRIVATE IS SILENT (*a silent desire*). Substituting the image schema in the *source domain* by a related color relationship predicts the semantic color associations SECRET IS DARK SATURATED BLUE/RED/BLACK and PRIVATE IS LOW BACKGROUND CONTRAST. These can be used to align the color design of the online banking website with the users' mental model of *security*, resulting in higher effectiveness, efficiency and user satisfaction, even across different cultures. Although the majority of the predictions of CMToC were empirically validated, further empirical studies that validate new metaphors are needed. Moreover, if designers want to achieve an even higher match with the users' mental models, multimodal instantiations of the image schema are recommended.

As color is a ubiquitous feature involved in merely all parts of HCI, the fields of application of CMToC are manifold. Compared to physical systems, color is a visual property that can be easily manipulated by a digital system (Hurtienne & Israel, 2013). When displaying information related to image-schematic dimensions through or supported by color, users can get a processing benefit if colors are used in a consistent way with respect to the natural correspondence with the image schema, and such links are understood by a variety of users. For example, a recent trend in UI design is to provide dimmable or inverted UIs that use modified light versions of the color scheme to be more easily visible in dark environments. Increasing the amount of dark colors in UIs can have an impact on how the semantic content is evaluated. According to CMToC, the color metaphor BRIGHT IS GOOD - BAD IS DARK predicts that a dark background color will bias users to judge the content more negative compared to a brighter background color. This prediction has been recently supported (Giron, 2016).

Taken together, such effects can be used to facilitate learning and intuitive use, increase safety, increase workplace productivity, but also to bias users in making certain judgements, like in persuasive technology. For example, temperature information is best represented through red (warm) and blue (cold), weight information should be preferably indicated by brighter (light) and darker colors (heavy). When evaluating temperature and weight, users will automatically consider such relevant color information, nudging their judgements into either direction. These relationships between color and image schemas also extend to more abstract content linked by the process of metaphorical mapping, at least when users are making conscious choices. Following these guidelines, designed information visualization artifacts can retain some of the experiential qualities that are typically associated with physical objects. But the application areas are not restricted to GUIs and displays. Advances in ubiquitous, tangible and IOT (internet of things) technology allow for online manipulation of an object's color in the users' environment. This promotes the application of color associations with other physical domains in fields like sensory marketing (Krishna, 2012), e.g. in sensorially enhanced glassware (Spence & Wan, 2015).

8.7.2. Applications Beyond HCI

Of course, semantic color associations can also be instrumentalized in contexts other than HCI. In his book 'Color & Light in Man Made Environment', Mahnke points to beneficial effects of applying the 'right' colors in our surrounding environment and warns of the negative impact a poorly designed environment can have on physical and psychological wellbeing (Mahnke & Mahnke, 1987). Examples of color usage in architecture, product design as well as marketing are mentioned, that also fall into the scope of the color associations promoted by CMToC. For example, Sunaga et al. recently investigated if package color and location influence how likely customers purchase products. The authors found that customers are more likely to choose light colored products (compared to dark colored products) when they are located in upper shelf positions (versus lower shelf positions), creating a coherent percept of vertical product placement and color-induced weight perception (Sunaga, Park, & Spence, 2016). Many similar effects just await empirical investigation as we are only at the beginning to advance our theoretical understanding of how color affects the way we think and act. CMToC as presented in this work provides researchers and practitioners alike with a sound starting point to guide further investigations in this field.

To conclude, although CMToC is a young field of investigation, its promises for predicting and explaining cross-cultural semantic color associations are stimulating, both for color psychology as well as UI design. This work has explored some relevant questions regarding its premises to derive implicit associations with physical properties and explicit associations with abstract concepts that can inform design decisions in a variety of UIs and pointed out open questions for future research. The results of the conducted studies show that the CMToC is worth pursuing as guidance for the application of color in UIs with the goal of developing a universal color design language and facilitating intuitive use with technology by building on humans' biological predisposition to treat color as a meaningful signal. The vast scope of application in both hard- and software UIs should encourage researchers and practitioners to carry on.

References

Abramov, I., Gordon, J., & Chan, H. (1991). Color appearance in the peripheral retina: effects of stimulus size. *Journal of the Optical Society of America. A, Optics and image science*, *8*(2), 404–414.

Adams, F. M., & Osgood, C. E. (1973). A Cross-Cultural Study of the Affective Meanings of Color. *Journal of Cross-Cultural Psychology*, *4*(2), 135–156.

Aires, B. (2016). Color and Sound : Physical and Psychophysical Relations and Color and Sound :. *Color Research & Application*, *19*(April 1994), 126–133.

Alexander, K. R. (1976). Influence of hue, value, and chroma on the perceived heaviness of colors. *Attention, Perception & Psychophysics*, *19*(1), 72–74.

Amano, S., & Kondo, T. (1999). *Nihongo-no Goi-Tokusei (Lexical properties of Japanese).* Sanseido Tokyo, Japan.

Anderson, M. L. (2003). Embodied Cognition: A field guide. *Artificial Intelligence*, *149*(1), 91–130.

Antle, A. N. (2011). Springboard: Designing Image Schema Based Embodied Interaction for an Abstract Domain. In *Whole body interaction* (pp. 101–119). Springer.

Asch, S. E. (1955). On the use of metaphor in the description of persons. *Clark University Conference on Expressive Language Behavior*, 29–38.

Aslam, M. M. (2006). Are You Selling the Right Colour? A Cross-cultural Review of Colour as a Marketing Cue. *Journal of Marketing Communications*, *12*(1), 15–30.

Bailey, R. J., Grimm, C. M., & Davoli, C. (2006). The real effect of warm-cool colors.

Balcer, C. A. (2014). Visual Cues Effects on Temperature Perception.

Baldauf, C. (1997). *Metapher und Kognition. Grundlagen einer neuen Theorie der Alltagsmetapher.* Peter Lang Verlag.

Ball, V. K. (1965). The Aesthetics of Color: A Review of Fifty Years of Experimentation. *Source: The Journal of Aesthetics and Art Criticism*, *23*(4), 441–452.

Ban, Y., Narumi, T., Fujii, T., Sakurai, S., Imura, J., Tanikawa, T., & Hirose, M. (2013). Augmented endurance: controlling fatigue while handling objects by affecting weight perception using augmented reality. In *Proceedings of chi 2013* (pp. 69–77).

Banerjee, P., Chatterjee, P., & Sinha, J. (2012). Is It Light or Dark? Recalling Moral Behavior Changes Perception of Brightness. *Psychological Science*, *23*(4), 407–409.

Banks, W. P., & Root, M. (1979). Semantic congruity effects in judgments of loudness. *Perception & Psychophysics*, *26*(2), 133–142.

Barchard, K. A., Grob, K. E., & Roe, M. J. (2016). Is sadness blue? The problem of using figurative language for emotions on psychological tests. *Behavior Research Methods*, 1–14.

Bargh, J. A., & Shalev, I. (2012). The substitutability of physical and social warmth in daily life. *Emotion*, *12*(1), 154–62.

Barker, R. G. (1978). *Habitats, environments, and human behavior: studies in ecological psychology and eco-behavioral science from the Midwest Psychological Field Station, 1947-1972*. San Francisco : Jossey-Baas.

Baron, R. A., Rea, M. S., & Daniels, S. G. (1992). Effects of indoor lighting (illuminance and spectral distribution) on the performance of cognitive tasks and interpersonal behaviors: The potential mediating role of positive affect. *Motivation and Emotion*, *16*(1), 1–33.

Barsalou, L. W. (2008). Grounded cognition. *Annual review of psychology*, *59*(August), 617–645.

Bayarri, S., Calvo, C., Costell, E., & Durán, L. (2001). Influence of Color on Perception of Sweetness and Fruit Flavor of Fruit Drinks. *Food Science and Technology International*, *7*(5), 399–404.

Becker, L. A. (2009). Effect size calculators. Retrieved from `http://www.uccs.edu/{~}lbecker/`

Bennett, C. A., & Rey, P. (1972). What's So Hot about Red? *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *14*(2), 149–154.

Bergum, B., & Bergum, J. (1981). Population stereotypes: An attempt to measure and define. In *Proceedings of the human factors society* (Vol. 25, pp. 662–665).

Berlin, B., & Kay, P. (1969). *Basic Color Terms: Their Universality and Evolution* (Vol. 19). Univ of California Press.

Berlyne, D. E. (1974). Aesthetics and psychobiology. *Curriculum in Art*, *4*(2/3), 205–211.

Berry, P. C. (1961). Effect of colored illumination upon perceived temperature. *Journal of Applied Psychology*, *45*(4), 248–250.

Bertelson, P., & Gelder, D. (2004). The psychology of multimodal perception. *Crossmodal space and crossmodal attention*, 141–177.

Bevan, W., & Dukes, W. F. (1953). Color as a variable in the judgment of size. *The American journal of psychology*, *66*(2), 283–288.

Birren, F. (2016). Color psychology and color therapy; a factual study of the influence of color on human life. *Academic Medicine*, *25*(4), 303.

Bonnardel, N., Piolat, A., & Le Bigot, L. (2011). The impact of colour on Website appeal and users' cognitive processes. *Displays*, *32*(2), 69–80.

Boroditsky, L. (2000). Metaphoric structuring: Understanding time through spatial metaphors. *Cognition*, *75*(1), 1–28.

Boroditsky, L., Fuhrman, O., & McCormick, K. (2011). Do English and Mandarin speakers think about time differently? *Cognition*, *118*(1), 126–132.

Boroditsky, L., & Prinz, J. (2008). What thoughts are made of. *Embodied Grounding: Social, Cognitive, Affective and Neuroscientific Approaches*, 98–115.

Brdar, M. (2001). *Metaphor in Cognitive Linguistics. Selected Papers from the 5th International Cognitive Linguistics Conference, Amsterdam, July 1997* (Vol. 11) (No. 3). John Benjamins Publishing.

Bretones, C. (2001). *Synaesthetic metaphors in English.* University of California at Berkeley and International Computer Science Institute, Berkeley.

Brogden, H. E., & Taylor, E. K. (1950). The theory and classification of criterion bias. *Educational and Psychological Measurement*, *10*(2), 159–183.

Brooks, R., & Brooks, R. (1991). Intelligence Without Representation. *Artificial Intelligence*, *47*(1-3), 139–159.

Brunswik, E. (1956). *Perception and the representative design of psy- chological experiments*. University of California Press.

Bubl, E., Kern, E., Ebert, D., Bach, M., & Tebartz Van Elst, L. (2010). Seeing gray when feeling blue? Depression can be measured in the eye of the diseased. *Biological Psychiatry*, *68*(2), 205–208.

Buckalew, L. W., & Coffield, K. E. (1982). An Investigation of Drug Expectancy as a Function of Capsule Color and Size and Preparation Form. *Journal of Clinical Psychopharmacology*, *2*(4), 245–248.

Buckingham, G., Cant, J. S., & Goodale, M. A. (2009). Living in a material world: how visual cues to material properties affect the way that we lift objects and perceive their weight. *Journal of neurophysiology*, *102*(6), 3111–3118.

Buechner, V. L., Maier, M. A., Lichtenfeld, S., & Elliot, A. J. (2015). Emotion expression and color: Their joint influence on perceived attractiveness and social position. *Current Psychology: A Journal for Diverse Perspectives on Diverse Psychological Issues*, *34*(2), 422–433.

Burton, G. J., & Moorhead, I. R. (1987). Color and spatial structure in natural scenes. *Applied Optics*, *26*(1), 157–170.

Caivano, J. L. (1998). Color and semiotics: A two-way street. *Color Research & Application*, *23*(6), 390–401.

Caldwell, J. A., & Jones, G. E. (1985). *The effects of exposure to red and blue light on physiological indices and time estimation* (Unpublished doctoral dissertation). University of Southern Mississippi.

Camgöz, N., Yener, C., & Güvenç, D. (2002). Effects of hue, saturation, and brightness on preference. *Color Research & Application*, *27*(3), 199–207.

Cant, J. S., Large, M. E., McCall, L., & Goodale, M. A. (2008). Independent processing of form, colour, and texture in object perception. *Perception*, *37*(1), 57–78.

Caporael, L. R. (1997). The Evolution of Trully Social Cognition: The Core Configurations Model. *Personality and Social Psychology Review*, *1*(4), 276–298.

Casasanto, D. (2008). Similarity and proximity: when does close in space mean close in mind? *Memory & Cognition*, *36*(6), 1047–1056.

Casasanto, D. (2009a). Embodiment of abstract concepts: good and bad in right- and left-handers. *Journal of experimental psychology. General*, *138*(3), 351–367.

Casasanto, D. (2009b). When is a linguistic metaphor a conceptual metaphor. *New directions in cognitive linguistics*, 127–145.

Casasanto, D. (2011). Different Bodies, Different Minds: The Body Specificity of Language and Thought. *Current Directions in Psychological Science*, *20*(6), 378–383.

Casasanto, D. (2014). Experiential origins of mental metaphors: Language, culture, and the body. *The Power of Metaphor: Examining its Influence on Social Life.*, 249–268.

Casasanto, D., & Bottini, R. (2014). Mirror Reading Can Reverse the Flow of Time. *Journal of Exjperimentai Psychology: General*, *143*(2), 473–479.

Casasanto, D., & Chrysikou, E. (2011). When Left Is "Right" - Motor Fluency Shapes Abstract Concepts. *Psychological Science*, *22*(4), 419–422.

Castanet, O. (1997). Pigmentary changes in aged and photoaged skin. *Archives of Dermatology*, *133*(10), 1296–1299.

Cavanagh, P., Tyler, C. W., & Favreau, O. E. (1984). Perceived velocity of moving chromatic gratings. *Journal of the Optical Society of America. A, Optics and image science*, *1*(8), 893–899.

Cavina-Pratesi, C., Kentridge, R. W., Heywood, C. A., & Milner, A. D. (2010). Separate channels for processing form, texture, and color: Evidence from fMRI adaptation and visual object agnosia. *Cerebral Cortex*, *20*(10), 2319–2332.

Centerbar, D. B., Schnall, S., Clore, G. L., & Garvin, E. D. (2008). Affective incoherence: when affective concepts and embodied reactions clash. *Journal of personality and social psychology*, *94*(4), 560–578.

Chadwick, A. C., & Kentridge, R. W. (2015). The perception of gloss: A review. *Vision Research*, *109*(April 2015), 221–235.

Chan, A. H. S., Shum, V. W. Y., Law, H. W., & Hui, I. K. (2003). Precise effects of control position, indicator type, and scale side on human performance. *International Journal of Advanced Manufacturing Technology*, *22*(5), 380–386.

Changizi, M. a., Zhang, Q., & Shimojo, S. (2006). Bare skin, blood and the evolution of primate colour vision. *Biology letters*, *2*(2), 217–221.

Chaparro, A., Stromeyer, C. F., Huang, E. P., Kronauer, R. E., & Eskew Jr, R. T. (1993). Colour is what the eye sees best. *Nature*, *361*(6410), 348–350.

Chen, C. R., & Ramaswamy, H. S. (2002). Color and Texture Change Kinetics in Ripening Bananas. *LWT - Food Science and Technology*, *35*(5), 415–419.

Cheng, K. M. (2002). *Quantitative evaluation of colour emotions* (Unpublished doctoral dissertation). The Hong Kong Polytechnic University.

Claessen, J. P., Overbeeke, C. J., & Smets, G. J. F. (1995). Puzzling colours. *Color Research & Application*, *20*(6), 388–396.

Clark, A. (1997). *Being there: Putting brain, body and word together again.* MIT press.

Clark, A. (2008). *Supersizing the mind: embodiment, action, and cognitive extension.* New York: Oxford University Press.

Clark, A., & Chalmers, D. (2010). The extended mind. *Analysis*, *58*(1), 7–19.

Clausner, T. C., & Croft, W. (1999). Domains and image schemas. *Cognitive Linguistics*, *10*(1), 1–31.

Cleveland, W. S. W. S., & McGill, R. (1983). A color-caused optical illusion on a statistical graph. *The American Statistician*, *37*(2), 101–105.

Clydesdale, F. (1993). Colour as a factor in food choice. *Critical reviews in Food Science and Nutrition*, *33*(1), 83–101.
Cohen, D., & Leung, A. K. Y. (2009). The hard embodiment of culture. *European Journal of Social Psychology*, *39*(7), 1278–1289.
Cohen, J. (1988). *Statistical power analysis: A computer program.* Routledge.
Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.
Collins, R. W. (2001). Software localization: Issues and methods. *Ecis*, *19*(2), 36–44.
Coules, J. (1955). Effects of photometric brightness on judgments of distance. *Journal of Experimental Psychology*, *50*(1), 19–25.
Courtney, A. J. (1986). Chinese Population Stereotypes: Color Associations. *Human Factors: The Journal of the Human Factors and Ergonomics*, *28*(1), 97–99.
Crisosto, C. H. (1994). Stone fruit maturity indices a descriptive review. *Postharvest News and Information*, *5*(6), 65N–68N.
Crozier, W. (1997). The psychology of colour preferences. *Surface Coatings International*, *80*(12), 577–585.
Cyr, D., Head, M., & Larios, H. (2010). Colur appeal in webside design within and across culturee: A multi-method evaluation. *International Journal of Human-Computer Studies*, *68*(1), 1–21.
Dael, N., Perseguers, M.-N., Marchand, C., Antonietti, J.-P., & Mohr, C. (2016). Put on that colour, it fits your emotion: Colour appropriateness as a function of expressed emotion. *Quarterly journal of experimental psychology*, *69*(8), 1619–1630.
Dahm, M. (2006). *Grundlagen der Mensch-Computer-Interaktion.* München: Pearson Studium.
Damasio, a. R. (1999). How the brain creates the mind. *Scientific American*, *281*(6), 112–117.
Davidoff, J. B. (1991). *Cognition through color.* Mit Press Cambridge, MA.
De Camp, J. E. (1917). The influence of color on apparent weight. A preliminary study. *Journal of experimental psychology*, *2*(5), 347.
Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences*, *10*(5), 204–211.
Deignan, A. (2005). *Metaphor and Corpus Linguistics.* John Benjamins Publishing.
Deignan, A. (2006). *Conceptual Metaphor Theory.* Retrieved 2016-12-11, from `http://creet.open.ac.uk/projects/metaphor-analysis/theories.cfm?paper=cmt`
Delwiche, J. (2004). The impact of perceptual interactions on perceived flavor. *Food Quality and Preference*, *15*(2), 137–146.
Derefeldt, G., Swartling, T., Berggrund, U., & Bodrogi, P. (2004). Cognitive Color. *Color Research and Application*, *29*(1), 7–19.
Deroy, O., & Spence, C. (2013). Why we are not all synesthetes (not even weakly so). *Psychonomic bulletin & review*, *20*(4), 643–64.

Deroy, O., & Valentin, D. (2011). Tasting liquid shapes: Investigating the sensory basis of cross-modal correspondences. *Chemosensory Perception*, *4*(3), 80–90.

Dhar, R., & Simonson, I. (2003). The effect of forced choice on choice. *Journal of Marketing Research*, *15*(8), 146–160.

Dhum, M., Oberfeld, D., Hecht, H., & Allendorf, U. (2006). Die Umgebungsfarbe bei der Weinverkostung beeinflusst globale Geschmacksurteile und maximalen Kaufpreis. In *48. tagung experimentell arbeitender psychologen, mainz.*

Dizdaroglu, M., Rao, G., Halliwell, B., & Gajewski, E. (1991). *Damage to the DNA bases in mammalian chromatin by hydrogen peroxide in the presence of ferric and cupric ions* (Vol. 285) (No. 2). Imprint Academic.

Dolscheid, S., Hunnius, S., Casasanto, D., & Majid, A. (2014). Prelinguistic Infants Are Sensitive to Space-Pitch Associations Found Across Cultures. *Psychological Science*, *25*(6), 1256–1261.

Dong, P., Huang, X., & Zhong, C.-B. (2015). Ray of Hope: Hopelessness Increases Preferences for Brighter Lighting. *Social Psychological and Personality Science*, *6*(1), 84–91.

Dougherty, R. F., Press, W. A., & Wandell, B. A. (1999). Perceived speed of colored stimuli. *Neuron*, *24*(4), 893–899.

Dresslar, F. (1894). Studies in the Psychology of Touch. *The American Journal of Psychology*, *6*(3), 313–368.

Drummond, P. D. (1997). Correlates of Facial Flushing and Pallor in Anger-Provoking Situations. *Personality and Individual Differences*, *23*(4), 575–582.

Duckitt, J., Wall, C., & Pokroy, B. (1999). Color bias and racial preference in White South African preschool children. *Journal of Genetic Psychology*, *160*(2), 143–154.

Duncker, K. (1939). The influence of past experience upon perceptual properties. *The American Journal of Psychology*, *52*(2), 255–265.

Dykes, J. R., & Cooper, R. G. (1978). An investigation of the perceptual basis of redundancy gain and orthogonal interference for integral dimensions. *Perception & Psychophysics*, *23*(1), 36–42.

Edlin, H. L. (1969). What Wood Is That - Manual of Wood Identification - Edlin,Hl.

Egusa, H. (1983). Effects of brightness, hue, and saturation on perceived depth between adjacent regions in the visual field. *Perception*, *12*(2), 167–175.

Ehrlich, S. B., Levine, S. C., & Goldin-Meadow, S. (2006). The importance of gesture in children's spatial reasoning. *Developmental psychology*, *42*(6), 1259–68.

Elliot, A. J. (2015). Color and psychological functioning: A review of theoretical and empirical work. *Frontiers in Psychology*.

Elliot, A. J., & Covington, M. (2001). Approach and Avoidance Motivation. *Educational Psychology Review*, *13*(2), 73–92.

Elliot, A. J., Eder, A. B., & Harmon-Jones, E. (2013). Approach-avoidance motivation and emotion: convergence and divergence. *Emotion Review*, *5*(3), 308–311.

Elliot, A. J., & Maier, M. A. (2007). Color and psychological functioning. *Current Directions in Psychological Science*, *16*(5), 250–254.

Elliot, A. J., & Maier, M. A. (2012a). Color-in-Context Theory. *Advances in Experimental Social Psychology*, *45*, 61–125.
Elliot, A. J., & Maier, M. A. (2012b). Color-in-Context Theory. *Advances in Experimental Social Psychology*, *45*, 66–125.
Elliot, A. J., & Maier, M. a. (2014). Color psychology: effects of perceiving color on psychological functioning in humans. *Annual review of psychology*, *65*, 95–120.
Elliot, A. J., Maier, M. A., Moller, A. C., Friedman, R., & Meinhardt, J. (2007). Color and psychological functioning: the effect of red on performance attainment. *Journal of Experimental Psychology: General*, *136*(1), 154–168.
Ernst, M. (2007). Learning to integrate arbitrary signals from vision and touch. *Journal of Vision*, *7*(5), 1–14.
Ernst, M., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433.
Evans, K. K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *Journal of Vision*, *10*(6).
Fairchild, M. D., & Fairchild, M. D. (2005). *Color Appearance Models*. John Wiley & Sons, Ltd.
Farne, M. (1977). Brightness as an indicator to distance: relative brightness per se or contrast with the background? *Perception*, *6*(3), 287–293.
Farrington, D. P., & Welsh, B. C. (2008). Effects of improved street lighting on crime: a systematic review. *Campbell Systematic Reviews*(13), 59.
Fastl, H. (2004). Audio-visual interactions in loudness evaluation. In *Proc. 18th int. congress on acoustics* (pp. 1161–1166).
Faubert, J. (1994). Seeing depth in colour: More than just what meets the eyes. *Vision Research*, *34*(9), 1165–1186.
Fay, A. J., & Maner, J. K. (2015). Embodied effects are moderated by situational cues: Warmth, threat, and the desire for affiliation. *British Journal of Social Psychology*, *54*(2), 291–305.
Fehrman, K., & Fehrman, C. (2000). *Color : The secret influence*. Prentice Hall.
Fenko, A., Otten, J. J., & Schifferstein, H. N. J. (2010). Describing product experience in different languages: The role of sensory modalities. *Journal of Pragmatics*, *42*(12), 3314–3327.
Fenko, A., Schifferstein, H. N. J., & Hekkert, P. (2009). Which senses dominate at different stages of product experience? *Undisciplined! Design Research Society Conference 2008*, 289/1–289/12.
Fetterman, A. K., Liu, T., & Robinson, M. D. (2015). Extending Color Psychology to the Personality Realm: Interpersonal Hostility Varies by Red Preferences and Perceptual Biases. *Journal of Personality*, *83*(1), 106–116.
Fetterman, A. K., Robinson, M. D., Gordon, R. D., & Elliot, A. J. (2011). Anger as Seeing Red: Perceptual Sources of Evidence. *Social Psychological and Personality Science*, *2*(3), 311–316.
Fetterman, A. K., Robinson, M. D., & Meier, B. P. (2012). Anger as "seeing red": Evidence for a perceptual association. *Cognition & Emotion*, *26*(June), 1–14.

Field, D. J. (1987). Relations Between the Statistics of Natural Images and the Response Properties of Cortical Cells. *Journal of the Optical Society of America*, *4*(12), 2379–2394.

Fiske, A. P. (2004). Four modes of constituting Relationships: Consubstantial Assimilation; Space, Magnitude, Time, and Force; Concrete Procedures; Abstract Symbolism. *Relational Models Theory. A contemporary Overview*, 61–146.

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83.

Folk, C. L. (2016). The role of color in the voluntary and involuntary guidance of selective attention. In A. J. Elliot, M. D. Fairchild, & A. Franklin (Eds.), *Handbook of color psychology* (pp. 481–501). Cambridge University Press.

Frank, M. G., & Gilovich, T. (1988). The dark side of self-perception and social-perception - black uniforms and aggression in professional sports. *Journal of Personality & Social Psychology*, *54*(1), 74–85.

Friedman, R. S., & Förster, J. (2010). Implicit affective cues and attentional tuning: An integrative review. *Psychological Bulletin*, *136*(5), 875–893.

Gao, X. P., Xin, J. H., Sato, T., Hansuebsai, A., Scalzo, M., Kajiwara, K., ... Billger, M. (2007). Analysis of cross-cultural color emotion. *Color Research and Application*, *32*(3), 223–229.

Garber, L. L., Hyatt, E. M., & Boya, Ö. Ö. (2008). The mediating effects of the appearance of nondurable consumer goods and their packaging on consumer behavior. *Product Experience*, 581–602.

Garber, L. L., Hyatt, E. M., & Nafees, L. (2015). The Effects of Food Color on Perceived Flavor: A Factorial Investigation in India. *Journal of Food Products Marketing*, *8*(4), 1–20.

Gegenfurtner, K. (2003). Cortical mechanisms of colour vision. *Nature Reviews. Neuroscience*, *4*(7), 563–572.

Gegenfurtner, K., & Rieger, J. (2000). Sensory and cognitive contribution of color to the recognition of natural scenes. *Current Biology*, *10*(13), 805–808.

Gegenfurtner, K., & Sharpe, L. (2001). *Color vision from genes to perception* (Vol. 26) (No. 6). Cambridge University Press.

Ghanem, L., Slobodenyuk, N., Elhajj, I., Kanso, A., & Jraissati, Y. (2015). Haptic and Tactile Adjectives Are Consistently Mapped onto Color Space. *Multisensory Research*, *29*(1-3), 253–278.

Gibbs, R. W. (1994). *The Poetics of Mind: Figurative Thought, Language, and Understanding* (Vol. 4) (No. Dec). Cambridge University Press.

Gibbs, R. W., Costa Lima, P. L., & Francozo, E. (2004). Metaphor is grounded in embodied experience. *Journal of Pragmatics*, *36*(7), 1189–1210.

Gil, S., & Bigot, L. L. (2015). Grounding context in face processing: Color, emotion, and gender. *Frontiers in Psychology*.

Gil, S., & Le Bigot, L. (2014). Seeing life through positive-tinted glasses: Color-meaning associations. *PLoS ONE*, *9*(8), e104291.

Gilbert, A. N., Martin, R., & Kemp, S. E. (1996). Cross-Modal Corresponsence between Vision and Olfraction: The Color of Smells. *The American Journal of Psychology*, *109*(3), 335–351.

Giron, L. (2016). *Black and White - Influence of Light and Dark Interfaces on Perceived Sentiment in Chats on the Example of Twitch.tv.*

Glenberg, A. M. (2010). Embodiment as a Unifying Perspective for Psychology. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(4), 586–596.

Goda, N., Tachibana, A., Okazawa, G., & Komatsu, H. (2014). Representation of the material properties of objects in the visual cortex of nonhuman primates. *The Journal of neuroscience*, *34*(7), 2660–2673.

Goldstein, K. (1942). Some experimental observations concerning the influence of colors on the function of the organism. *Occupational Therapy*, *21*(3), 147–151.

Gorn, G. J., Chattopadhyay, A., Sengupta, J., & Tripathi, S. (2004). Waiting for the Web: How Screen Color Affects Time Perception. *Journal of Marketing Research*, *41*(2), 215–225.

Gottfried, J. A., & Dolan, R. J. (2003). The nose smells what the eye sees: Crossmodal visual facilitation of human olfactory perception. *Neuron*, *39*(2), 375–386.

Grady, J. (1997). *Foundations of meaning: Primary metaphors and primary scenes.* University of California, Berkeley.

Grady, J. (1999). A typology of motivation for metaphor: Correlations vs. resemblances. *Metaphor in cognitive linguistics*, 79–100.

Grady, J. (2005). Primary metaphors as input to conceptual metaphors. *Journal of Pragmatics*, *37*(10), 1595–1614.

Green, W., Hasson, S., Mohammed, S., Phillips, C., Richards, P., Smith ME, & White, A. (1982). Effect of viewing selected colors on the performance of gross and fine motor tasks. *Perceptual and Motor Skills*, *54*(3), 778.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (2008). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480.

Guéguen, N., & Jacob, C. (2014). Coffee cup color and evaluation of a beverage's "warmth quality". *Color Research and Application*, *39*(1), 79–81.

Gundlach, C., & Macoubrey, C. (1931). The effect of color on apparent size. *The American Journal of Psychology*, *43*(1), 109–111.

Häberle, C. J. (2000). *Farben in Europa* (Unpublished doctoral dissertation). Universität Duisburg-Essen.

Hager, W., & Hasselhorn, M. (1994). Handbuch deutschsprachiger Wortnormen.

Hahn, A. C., Whitehead, R. D., Albrecht, M., Lefevre, C. E., & Perrett, D. I. (2012). Hot or not? Thermal reactions to social contact. *Biology Letters*, *8*(5), 864–867.

Hampe, B. (2005a). *Image Schemas in Cognitive Linguistics: Introduction.* Walter de Gruyter.

Hampe, B. (2005b). On the role of iconic motivation in conceptual metaphor - Has metaphor theory come full circle? *Outside-In Inside-Out*, 39–66.

Hampe, B. (2005c). When down is not bad, and up not good enough: A usage-based assessment of the plus–minus parameter in image-schema theory. *Cognitive Linguistics*, *16*(2005), 81–112.

Hanbury, A., & Serra, J. (2003). Circular statistics applied to colour images. In *8th computer vision winter workshop* (Vol. 91, pp. 53–71).

Harnad, S. (1990). The Symbol Grounding Problem. *Physica D: Nonlinear Phenomena*, *42*(1), 335–346.

Harrar, V., Piqueras-Fiszman, B., & Spence, C. (2011). There's more to taste in a coloured bowl. *Perception*, *40*(7), 880–882.

Hassenzahl, M., Heidecker, S., Eckoldt, K., Diefenbach, S., & Hillmann, U. (2012, dec). All You Need is Love. *ACM Transactions on Computer-Human Interaction*, *19*(4), 1–19.

Hatta, T., Yoshida, H., Kawakami, A., & Okamoto, M. (2002). Color of computer display frame in work performance, mood, and physiological response. *Perceptual and Motor Skills*, *94*(1), 39–46.

Heise, D. (1970). The semantic differential and attitude research. *Attitude measurement*(4), 235–253.

Heller, E. (1999). *Wie farben wirken. Farbpsychologie, Farbsymbolik, Kreative Farbgestaltung.* Rowohlt.

Hemphill, M. (1996). A Note on Adults' Color–Emotion Associations. *The Journal of Genetic Psychology*, *157*(3), 275–280.

Higgins, E. T., Horwitz, E. K., Horwitz, M. B., Cope, J., Horwitz, E. K., Horwitz, M. B., & Cope, J. (1987). Self-discrepancy: a theory relating self and affect. *Psychological review*, *94*(3), 319–340.

Hill, R. a., & Barton, R. a. (2005). Psychology: red enhances human performance in contests. *Nature*, *435*(May), 293.

Hinkle D, Jurs S, & Wiersma W. (2003). *Applied Statistics for the Behavioral Sciences.* Houghton Mifflin, Boston.

Ho, H. (2015). Color-temperature correspondence and its impact on object temperature perception. *Proceedings of VINCI 2015*, *13*.

Ho, H., Iwai, D., Yoshikawa, Y., Watanabe, J., & Nishida, S. (2014). Combining colour and temperature: A blue object is more likely to be judged as warm than a red object. *Scientific reports*, *4*, 5527.

Ho, H., Van Doorn, G., Kawabe, T., Watanabe, J., & Spence, C. (2014). Colour-temperature correspondences: When reactions to thermal stimuli are influenced by colour. *PLoS ONE*, *9*(3), e91854.

Hofstede, G. (2001). *Culture's consequences: comparing values, behavior, institutions and organizations across nations.* Sage Publications, Inc.

Howard, C. M., & Burnidge, J. A. (1994). Colors in natural landscapes. *Journal of the Society for Information Display*, *2*(1), 47–55.

Hulse, F. S. (1967). Selection for skin color among the Japanese. *American Journal of Physical Anthropology*, *27*(2), 143–155.

Humphrey, N. (1976). The colour currency of nature. *Colour for architecture*, 95–98.

Hupka, R. B., Zaleski, Z., Otto, J., Reidl, L., & Tarabrina, N. V. (1997). The colors of anger, envy, fear, and jealousy: A cross-cultural study. *Journal of cross-cultural psychology*, *28*(2), 156–171.

Hurtienne, J. (2011). *Image Schemas and Design for Intuitive Use.*

Hurtienne, J., & Blessing, L. (2007). Design for Intuitive Use-Testing image schema theory for user interface design. In *16th international conference on engineering design* (pp. 1–12).

Hurtienne, J., & Israel, J. H. (2013). PIBA-DIBA or How to Blend the Digital with the Physical. In *Workshop blended interaction: Envisioning future collaborative interactive spaces at acm sigchi conference on human factors in computing systems.*

Hurtienne, J., & Langdon, P. (2009). Prior knowledge in inclusive design: The older, the more intuitive. In *Proceedings of the 23rd british computer society human computer interaction workshop and conference (hci'09).*

Hurtienne, J., & Meschke, O. (2016). Soft Pillows and the Near and Dear: Physical-to-Abstract Mappings with Image-Schematic Metaphors. In *Proceedings of the tei'16: Tenth international conference on tangible, embedded, and embodied interaction* (pp. 324–331).

Hurtienne, J., Stößel, C., & Weber, K. (2009). Sad is heavy and happy is light: population stereotypes of tangible object attributes. In *Tangible and embedded interaction* (pp. 61–68).

Huskisson, E. C. (1974). Simple analgesics for arthritis. *British medical journal*, *4*(5938), 196–200.

IJzerman, H., & Koole, S. L. (2011). From perceptual rags to metaphoric riches-bodily, social, and cultural constraints on sociocognitive metaphors: comment on Landau, Meier, and Keefer (2010). *Psychological Bulletin*, *137*(2), 355–361.

IJzerman, H., & Semin, G. R. (2010). Temperature perceptions as a ground for social proximity. *Journal of Experimental Social Psychology*, *46*(6), 867–873.

Ishihara, S. (1917). *Tests for Colour Blindness.* Handaya Hongo Harukich, Tokyo.

Ishii, H. (2008). Tangible bits: beyond pixels. In *Proceedings of the 2nd international conference on tangible and embedded intreaction* (pp. xv–xxv).

Jacobs, K. W., & Hustmyer Jr, F. E. (1974). Effects of Four Psychological Primary Colors on GSR, Heart Rate and Respiration Rate. *Perceptual and Motor Skills*, *38*(June), 763–766.

Jahanian, A., Vishwanathan, S. V. N., & Allebach, J. P. (2015). Colors - Messengers of Concepts : Visual Design Mining for Learning Color Semantics.

Jäkel, O. (2003). *Wie Metaphern Wissen schaffen.* Hamburg: Dr. Kovac.

Jalil, N. A., Yunus, R. M., & Said, N. S. (2012). Environmental Colour Impact upon Human Behaviour: A Review. *Procedia - Social and Behavioral Sciences*, *35*, 54–62.

Jian, W. S. X. (2002). Synaesthesia and Conceptual Metaphor. *Foreign Language Research*, *3*.

Johns, E. H., & Sumner, F. C. (1948). Relation of the brightness differences of color to their apparent distances. *The Journal of Psychology*, *26*(1), 25–29.

Johnson, M. (1987). *The body in the mind. The bodily basis of meaning, imagination, and reason.* University of Chicago Press Chicago.

Kaiser, P. K. (1984). Physiological Response to Color : A Critical Review. *Color Research & Application*, *9*(1), 29–36.

Kaplan, P. D., Polefka, T., Grove, G., Daly, S., Jumbelic, L., Harper, D., ... Bianchini, R. (2011). Grey hair: Clinical investigation into changes in hair fibres with loss of pigmentation in a photoprotected population. *International Journal of Cosmetic Science*, *33*(2), 171–182.

Karana, E., & Hekkert, P. (2010). User-Material-Product interrelationships in attributing meanings. *International Journal of Design*, *4*(3), 43–52.

Karana, E., Hekkert, P., & Kandachar, P. (2007). Sensorial properties of materials for creating expressive meanings. In *1st kansei engineering and emotion ...* (pp. 7–11).

Karana, E., Hekkert, P., & Kandachar, P. (2009). Meanings of materials through sensorial properties and manufacturing processes. *Materials and Design*, *30*(7), 2778–2784.

Kaufmann, C. (2006). *Zur Semantik der Farbadjektive rosa, pink und rot* (Unpublished doctoral dissertation). LMU München.

Kemp, S. E., & Gilbert, A. N. (1997). Odor intensity and color lightness are correlated sensory dimensions. *American Journal of Psychology*, *110*(1), 35–46.

Kereney, G. E. (1966). Hue Preferences as a Function of Ambient Temperatures. *Australian Journal of Psychology*, *18*(3), 271–275.

Kim, J., Marlow, P. J., & Anderson, B. L. (2012). The dark side of gloss. *Nature Neuroscience*, *15*(11), 1590–1595.

Kim, Y. (2013). Can eyes smell? cross-modal correspondences between color hue-tone and fragrance family. *Color Research and Application*, *38*(2), 139–156.

Klein, G. S. (1964). Semantic Power Measured Through the Interference of Words With Color-Naming. *The American Journal of Psychology*, *77*(4), 576–588.

Koch, C., & Koch, E. C. (2003). Preconceptions of taste based on color. *The Journal of Psychology*, *137*(3), 233–242.

Kort, Y. A. W. D., Ijsselsteijn, W. A., Haans, A., Lakens, D., & Kalinauskaite, I. (2010). De-escalate : Defusing escalating behaviour through the use of interactive light scenarios. *Human-Technology Interaction, School of Innovative Sciences, & Intelligent Lighting Institute (the Netherlands)*, 1–4.

Kövecses, Z. (1990). *Emotion Concepts.* New York: Springer.

Kövecses, Z. (1997). Metaphor: Does it constitute or reflect cultural models? *Metaphor in Cognitive Linguistics*, 167–188.

Kövecses, Z. (2002). *Metaphor. A Practical Introduction.* Oxford University Press.

Kövecses, Z. (2005). *Metaphor in Culture: Universality and Variation.* Cambridge University Press.

Kövecses, Z. (2006). Language, mind, and culture.

Krantz, J. H. (2012). *Experiencing Sensation and Perception.* Pearson Education (Us).

Kremers, J., Silveira L C L, Yamada, E. S., & Lee, B. B. (1999). The ecology and evolution of primate color vision. *Color Vision: From Genes to Perception*, 113–122.

Krenn, B. (2014). The impact of uniform color on judging tackles in association football. *Psychology of Sport and Exercise*, *15*(2), 222–225.

Krishna, A. (2012). An integrative review of sensory marketing: Engaging the senses to affect perception, judgment and behavior. *Journal of Consumer Psychology*, *22*(3), 332–351.

Kutas, G., Gócza, K., Bodrogi, P., & Schanda, J. (2004). Colour size effect. In *Conference on colour in graphics, imaging, and vision* (pp. 887–890).

Kwallek, N., & Lewis, C. M. (1996). Effects of environmental colour on males and females: a red or white or green office. *Applied Ergonomics*, *177*(4), 275–278.

Labrecque, L. I., Patrick, V. M., & Milne, G. R. (2010). The Marketers' Prismatic Palette: A Review of Color Research and Future Directions. *Psychology and Marketing*, *30*(2), 187–202.

Lajos, J., & Paris, H. E. C. (2009). Effects of Color on Consumers' Perceptions of Package Volumes. *Journal of Contingencies and Crisis Management*, *12*(1994), 838–840.

Lakens, D. (2012). Polarity correspondence in metaphor congruency effects: structural overlap predicts categorization times for bipolar concepts presented in vertical space. *Journal of experimental psychology. Learning, memory, and cognition*, *38*(3), 726–36.

Lakens, D. (2014). Grounding Social Embodiment. *Social Cognition*, *32*(Supplement), 168–183.

Lakens, D., Fockenberg, D. a., Lemmens, K. P. H., Ham, J. R., & Midden, C. J. (2013, nov). Brightness differences influence the evaluation of affective pictures. *Cognition & emotion*, *27*(7), 1225–1246.

Lakens, D., Semin, G. R., & Foroni, F. (2012). But for the bad, there would not be good: Grounding valence in brightness through shared relational structures. *Journal of Experimental Psychology: General*, *141*(3), 584–594.

Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. University of Chicago Press.

Lakoff, G. (1990). The invariance hpothesis: Is abstract reason based on image-schemes? *Cognitive Linguistics*, *1*(1990), 39–74.

Lakoff, G. (1993). The Contemporary Theory of Metaphor. *Metaphor and Thought*.

Lakoff, G. (1994). *Master Metaphor List*. University of California.

Lakoff, G., & Johnson, M. (1980). Metaphors we live by.

Lakoff, G., & Johnson, M. (1999). *Philosophy In The Flesh: The Embodied Mind And Its Challenge To Western Thought*. New York: Basic books.

Lakoff, G., & Turner, M. (2009). *More Than Cool Reason - a Field Guide to Poetic Metaphor*. University of Chicago Press.

Lalanne, C., & Lorenceau, J. (2004). Crossmodal integration for perception and action. *Journal of Physiology Paris*, *98*(1), 265–279.

Landau, M. J., Keefer, L. A., & Meier, B. P. (2011). Wringing the Perceptual Rags: Reply to IJzerman and Koole (2011). *Psychological Bulletin*, *137*(2), 362–365.

Landau, M. J., Meier, B. P., & Keefer, L. A. (2010). A Metaphor-Enriched Social Cognition. *Psychological Bulletin*, *136*(6), 1045–1067.

Landis, J. R., & Koch, G. G. (1977, mar). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

Langdon, P., Clarkson, J., & Robinson, P. (2008). *Designing Inclusive Futures* (P. Langdon, J. Clarkson, & P. Robinson, Eds.). London: Springer London.

Lebrecht, S. (2012). "Micro-valences": Affective valence in "neutral" everyday objects.

Lederman, S. J., Thorne, G., & Jones, B. (1986). Perception of texture by vision and touch: multidimensionality and intersensory integration. *Journal of Experimental Psychology. Human Perception and Performance*, *12*(2), 169–180.

Lee, D. (2001). *Cognitive Linguistics: An Introduction.* Lawrence Erlbaum Associates Publishers.

Legge, G. E., & Chung. (2010). *Encyclopedia of Perception.* Sage Publications.

Levitan, C. A., Zampini, M., Li, R., & Spence, C. (2008). Assessing the role of color cues and people's beliefs about color-flavor associations on the discrimination of the flavor of sugar-coated chocolates. *Chemical Senses*, *33*(5), 415–423.

Levy, B. I. (1984). Research into the psychological meaning of color. *American Journal of Art Therapy*, *23*(2), 58–62.

Lewallen, K. S., & Marini, R. P. (2003). Relationship between Flesh Firmness and Ground Color in Peach as Influenced by Light and Canopy Position. *Journal of American Society Horticultural Science*, *128*(2), 163–170.

Lewin, K. (1951). Field theory in social science.

Lewis, D. E., Pearson, J., & Khuu, S. K. (2013). The Color "Fruit": Object Memories Defined by Color. *PLoS ONE*, *8*(5), e64960.

Li, H., Sun, X., & Zhang, K. (2007). Culture-centered design: Cultural factors in interface usability and usability tests. In *Eighth acis international conference on software engineering, artificial intelligence, networking, and parallel/distributed computing* (Vol. 3, pp. 1084–1088).

Lindgaard, G., Dudek, C., Sen, D., Sumegi, L., & Noonan, P. (2011). An Exploration of Relations Between Visual Appeal, Trustworthiness and Perceived Usability of Homepages. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *18*(1), 1–30.

Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *7*(11), 3416–3468.

Löffler, D. (2014a). Happy is Pink: Designing for Intuitive Use with Color-to-abstract Mappings. In *Chi '14 extended abstracts on human factors in computing systems* (pp. 323–326).

Löffler, D. (2014b). Population stereotypes of color attributes for tangible interaction design. In *8th international conference on tangible, embedded and embodied interaction - tei'14* (pp. 285–288).

Löffler, D., Arlt, L., Toriizuka, T., Tscharn, R., & Hurtienne, J. (2016). Substituting Color for Haptic Attributes in Conceptual Metaphors for Tangible Interaction Design. In *Tei'16: Tenth international conference on tangible, embedded, and embodied interaction* (pp. 118–125). ACM.

Löffler, D., Paier, W., Toriizuka, T., Ikeda, M., & Hurtienne, J. (2015). Augmented Happiness: Simple Color Changes Influence Users' Conceptual Choices. In *Human-computer interaction* (pp. 248–255).

Lorentz, E., Ekstrand, C., Gould, L., & Borowsky, R. (2016). Red-hot: How colour and semantic temperature processing interact in a Stroop-like paradigm. *Visual Cognition*, *24*(2), 173–181.

Lucassen, M. P., Gevers, T., & Gijsenij, A. (2011). Texture affects color emotion. *Color Research and Application*, *36*(6), 426–436.

Ludden, G., & Kudrowitz, B. (2012). Surprise and humor in product design. *Humor*, *3*(2012), 285–309.

Ludden, G., Schifferstein, H., & Hekkert, P. (2009). Visual-Tactual Incongruities in Products as Sources of Surprise. *Empirical Studies of the Arts*, *27*(1), 61–87.

Ludden, G., Schifferstein, H., & Hekkert, P. (2012). Beyond surprise: A longitudinal study on the experience of visual-tactual incongruities in products. *International Journal of Design*, *6*(1), 1–10.

Ludwig, V. U., & Simner, J. (2013). What colour does that feel? Tactile-visual mapping and the development of cross-modality. *Cortex*, *49*(4), 1089–1099.

Lynch, D., & Livingston, W. (2001). *Colour and Light in Nature*. Cambridge University Press.

Macaranas, A., Antle, A. N., & Riecke, B. E. (2012). Bridging the Gap: Attribute and Spatial Metaphors for Tangible Interface Design. In *Proceedings of the sixth international conference on tangible, embedded and embodied interaction - tei '12* (pp. 161–168).

MacDonald, L. W. (1999). Using color effectively in computer graphics. *IEEE Computer Graphics and Applications*, *19*(4), 20–35.

Madden, J.T., Hewett, K., Roth, M. (2000). Managing Images in Different Cultures: A Cross-National Study of Colour Meaning and Preferences. *Journal of International Marketing*, *8*(4), 90–107.

Maga, J. A. (1974). Influence of color on taste thresholds. *Chemical Senses*, *1*(1), 115–119.

Mahnke, F. (1996). Color, environment and human response. *Chapter 8: Biological effects of light*, 102–122.

Mahnke, F., & Mahnke, R. (1987). Color and Light in Man Made Environments.

Mandler, J. (1992). How to build a baby: Conceptual primatives. *Psychological Review*, *99*(4), 589–604.

Mandler, J. (2008). On the birth and growth of concepts. *Philosophical psychology*, *21*(2), 207–230.

Marks, L. E. (1990). Synaesthesia: Perception and Metaphor. *Aesthetic Illusion: Theoretical and Historical Approaches*, 28–40.

Marr, D. (1982). *Vision.* San Francisco: Freeman.

McElhanon, K. A. (2006). From Simple Metaphors to Conceptual Blending: The Mapping of Analogical Concepts and the Praxis of Translation. *Journal of translation*, *2*(1), 31–81.

McLeod, C. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*(2), 163–203.

McManus, C., Jones, A. L., & Cottrell, J. (1981). The Aesthetics of Color. *Perception*, *10*(6), 651–666.

Mead, M., & Baldwin, J. (2001). *A Rap on Race* (Vol. 28). Laurel.

Mehta, R. P., & Zhu, R. J. (2009). Blue or red? Exploring the effect of color on cognitive task performances. *Science*, *323*(5918), 1226–1229.

Meier, B. P. (2016). *Do metaphors color our perception of social life?* (A. J. Elliot & M. D. Fairchild, Eds.). Cambridge, U.K.: Cambridge University Press.

Meier, B. P., D'Agostino, P. R., Elliot, A. J., Maier, M. A., & Wilkowski, B. M. (2012). Color in context: Psychological context moderates the influence of red on approach- and avoidance-motivated behavior. *PLoS ONE*, *7*(7), e40333.

Meier, B. P., Fetterman, A. K., & Robinson, M. D. (2015). Black and White as Valence Cues. *Social Psychology*, *46*(3), 174–178.

Meier, B. P., Robinson, M. D., & Caven, A. J. (2008). Why a Big Mac Is a Good Mac: Associations between Affect and Size. *Basic and Applied Social Psychology*, *30*(1), 46–55.

Meier, B. P., Robinson, M. D., & Clore, G. L. (2004). Why good guys wear white: Automatic inferences about stimulus valence based on brightness. *Psychological Science*, *15*(2), 82–87.

Meier, B. P., Robinson, M. D., Crawford, L. E., & Ahlvers, W. J. (2007). When "Light" and "Dark" Thoughts Become Light and Dark Responses: Affect Biases Brightness Judgments. *Emotion*, *7*(2), 366–376.

Meier, B. P., Schnall, S., Schwarz, N., & Bargh, J. A. (2012). Embodiment in Social Psychology. *Topics in Cognitive Science*, *4*(4), 705–716.

Melara, R. D. (1989). Dimensional interaction between color and pitch. *Journal of experimental psychology. Human perception and performance*, *15*(1), 69–79.

Menzel, D., Fastl, H., Graf, R., & Hellbrück, J. (2008). Influence of vehicle color on loudness judgments. *The Journal of the Acoustical Society of America*, *123*(5), 2477–2479.

Michael, G. A., Galich, H., Relland, S., & Prud'hon, S. (2010). Hot colors: The nature and specificity of color-induced nasal thermal sensations. *Behavioural Brain Research*, *207*(2), 418–428.

Moos, A., Simmons, D., Simner, J., & Smith, R. (2013). Color and texture associations in voice-induced synesthesia. *Frontiers in Psychology*, *4*(568).

Morgan, G. A., Goodson, F. E., & Jones, T. (1975). Age differences in the associations between felt temperatures and color choices. *The American journal of psychology*, *88*(1), 125–130.

Morrison, C. T., Oates, T., & King, G. (2001). *Grounding the Unobservable in the Observable: The Role and Representation of Hidden State in Concept Formation and Refinement* (Tech. Rep.). DTIC Document.

Moseley, G. L., & Arntz, A. (2007). The context of a noxious stimulus affects the pain it evokes. *Pain*, *133*(1), 64–71.

Nakagawa, S. (2004). A farewell to Bonferroni: the problem of low statistical power and publication bias. *Behavioral Ecology*, *15*(6), 1044–1045.

Nakamura, T., Hoshino, H., Sato, T., & Kajiwara, K. (1996). Attempt for quantitative evaluation of cool-warm color image. *Sen-I Gakkaishi*, *52*(1), 27–31.

Nakano, M., Tanabe, S., Mori, Y., Ikegami, B., & Fujita, I. (2005). Expansive and contractive size perception with color patches. *Journal of Vision*, *5*(8), 1027.

Naumann, A., Hurtienne, J., Israel, J. H., Mohs, C., Kindsmüller, M. C., Meyer, H. A., ... IUUI Research Group (2007). Intuitive use of user interfaces: Defining a vague concept. In *Engineering psychology and cognitive ergonomics* (pp. 128–136).

Naz, K., & Epps, H. (2004). Relationship between color and emotion: a study of college students. *College Student J*, *38*(3), 396–405.

Newton, I. (1730). Opticks: or, A treatise of the reflexions, refractions, inflexions and colours of light. *Optics*, 1642–1727.

Nguyen-Tri, D., Overbury, O., & Faubert, J. (2003). The role of lenticular senescence in age-related color vision changes. *Investigative Ophthalmology and Visual Science*, *44*(8), 3698–3704.

Nisbett, R. E. (2003). *The Geography of Thought: How Asians and Westerners Think Differently... and Why.* Simon and Schuster.

Nisbett, R. E., Peng, K., Choi, I., Ames, D., Atran, S., Cheng, P., ... Yates, F. (2001). Culture and Systems of Thought: Holistic Versus Analytic Cognition. *Psychological Review*, *108*(2), 291–310.

Nunnally, J. C., & Bernstein, I. H. (1978). *Psychometric theory.* New York: McGraw-Hill.

Oberfeld, D., Hecht, H., Allendorf, U., & Wickelmaier, F. (2009). Ambient Lighting Modifies the Flavor of Wine. *Journal of Sensory Studies*, *24*(6), 797–832.

Okubo, M., & Ishikawa, K. (2011). Automatic semantic association between emotional valence and brightness in the right hemisphere. *Cognition & Emotion*, *25*(7), 1273–1280.

Oliva, A., & Schyns, P. G. (2000). Diagnostic Colors Mediate Scene Recognition. *Cognitive Psychology*, *41*(2), 176–210.

Onasanya, S. (2002). The Effect of Colour on Students' Cognitive Performance in Instruction Using Photographic Prints. *Ilorin Journal of Education*, *21*, 156–166.

Onofrei, M., Hunt, J., Siemienczuk, J., Touchette, D. R., & Middleton, B. (2004). A first step towards translating evidence into practice: Heart failure in a community practice-based research network. *Informatics in Primary Care*, *12*(3), 139–145.

O'Regan, K., & Noe, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, *24*(5), 939–1031.

Ortiz, M. J. (2011). Primary metaphors and monomodal visual metaphors. *Journal of Pragmatics*, *43*(6), 1568–1580.

Ortony, A. (1979). Beyond literal similarity. *Psychological Review*, *86*(3), 161–180.

Osgood, C. E. (1969). The Nature and Measurement of Meaning. *Origin and Nature of the Semantic Differential Technique - a sourcebook*, *49*(3), 3–41.

Osgood, C. E., May, W. H., & Miron, M. S. (1975). Cross-Cultural Universals of Affective Meaning.

Österbauer, R. A., Matthews, P. M., Jenkinson, M., Beckmann, C. F., Hansen, P. C., & Calvert, G. A. (2005). Color of Scents: Chromatic Stimuli Modulate Odor Responses in the Human Brain. *Journal of Neurophysiology*, *93*(6), 3434–3441.

Ou, L.-C., Luo, M., Sun, P.-L., Hu, N.-C., Chen, H.-S., Guan, S.-S., ... Richter, K. (2012). A cross-cultural comparison of colour emotion for two colour-combinations. *Color Research & Application*, *37*(1), 23–43.

Ou, L. C., Luo, M. R., Woodcock, A., & Wright, A. (2004). A study of colour emotion and colour preference. Part III: Colour preference modeling. *Color Research and Application*, *29*(5), 381–389.

Oyama, T., Tanaka, Y., & Chiba, Y. (1962). Affective dimensions of colors: A cross-cultural study. *Japanese Psychological Research*, *4*(2), 78–91.

Palmer, S. E., & Schloss, K. B. (2009). An ecological valence theory of human color preference. *Journal of Vision*, *9*(8), 2663–2668.

Palmer, S. E., Schloss, K. B., Xu, Z., & Prado-León, L. R. (2013). Music-color associations are mediated by emotion. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(22), 8836–8841.

Patsouras, C., Filippou, T., & Fastl, H. (2002). Influences of color on the loudness judgement. In *Forum acusticum* (Vol. 2, pp. 1–6).

Payne, M. C. (1964). Color As an Independent Variable in Perceptual Research. *Psychological Bulletin*, *61*(3), 199–208.

Pengnate, S., & Antonenko, P. (2013). A Multimethod Evaluation of Online Trust and Its Interaction with Metacognitive Awareness: An Emotional Design Perspective. *International Journal of Human-Computer Interaction*, *29*(9), 582–593.

Pinkerton, E., & Humphrey, N. K. (1974). The apparent heaviness of colours. *Nature*, *250*(5462), 164–165.

Piqueras-Fiszman, B., Alcaide, J., Roura, E., & Spence, C. (2012). Is it the plate or is it the food? Assessing the influence of the color (black or white) and shape of the plate on the perception of the food placed on it. *Food Quality and Preference*, *24*(1), 205–208.

Piqueras-Fiszman, B., Velasco, C., & Spence, C. (2012). Exploring implicit and explicit crossmodal colour-flavour correspondences in product packaging. *Food Quality and Preference*, *25*(2), 148–155.

Plack, J. J., & Shick, J. (1976). Elementary school child's perception of weight as a function of color. *Journal of the Association for the Study of Perception*, *11*(2),

16–20.
Posner, M. I., & Snyder, C. R. R. (1975). Facilitation and inhibition in the processing of signals. *Attention and performance V*, 669–682.
Rader, T., Morigana, M., Matsui, T., Fastl, H., Kuwano, S., & Namba, S. (2004). Crosscultural effects in audio-visual interactions. In *In transactions of the technical committee of noise and vibration of the acoustical society of japan* (pp. 1–6).
Raskin, J. (1994). Viewpoint: Intuitive equals familiar. *Communications of the ACM*, *37*(9), 17–18.
Reece, A. G., & Danforth, C. M. (2016). Instagram photos reveal predictive markers of depression. *arXiv:1608.03282*. Retrieved from `http://arxiv.org/abs/1608.03282`
Regier, T., Kay, P., & Cook, R. S. (2005). Focal colors are universal after all. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(23), 8386–8391.
Rentschler, I., Jüttner, M., Unzicker, A., & Landis, T. (1999). Innate and learned components of human visual preference. *Current Biology*, *9*(13), 665–671.
Rheingans, P. (1997). Color perception and applications. *Siggraph'97. Course*. Retrieved from `http://www.csee.umbc.edu/{~}rheingan/SIGGRAPH/color.intro.pdf`
Richter, T., & Zwaan, R. A. (2009). Processing of colour words activates colour representations. *Cognition*, *111*(3), 383–389.
Roberson, D., Davidoff, J., Davies, I. R. L., & Shapiro, L. R. (2005). Color categories: Evidence for the cultural relativity hypothesis. *Cognitive Psychology*, *50*(4), 378–411.
Robinson, E. J. (1954). The influence of photometric brightness on judgments of size. *The American journal of psychology*, *67*(3), 464–474.
Robinson, M. D. (2007). Lives lived in milliseconds. *Handbook of Research Methods in Social and Personality Psychology 2*, 345–369.
Rock, I., & Victor, J. (1964). Vision and touch: An experimentally created conflict between the two senses. *Science*, *143*(3606), 594–596.
Rolls, E. T. (2000). On the brain and emotion. *The Behavioral and brain sciences*, *23*(2), 177–191.
Roullet, B., & Droulers, O. (2005). Pharmaceutical packaging color and drug expectancy. *Advances in Consumer Research*, *32*, 164–171.
Sable, P., & Akcay, O. (2010). Color: Cross Cultural Marketing Perpective as to What Governs our Response to it. *Asbbs*, *17*(1), 950–954.
Saiz-Jimenez, C. (2004). *Air pollution and cultural heritage* (Vol. 1). CRC Press.
Sambre, P. (2000). Metaphor and Metonymy at the Crossroads: A Cognitive Perspective. *Pragmatics*, *10*(3), 344–345.
Sato, T., Kajiwara, K., Hoshino, H., & Nakamura, T. (1997a). The attempt of quantitative assessment for "vivid-sombre", "gaudy-plain", "striking-subdued", "dynamic-passive", "distinct-vague", "transparent-turbid", "soft-hard" and "strong-weak" colour images. *Sen-I Gakkaishi*, *53*(4), 131–138.

Sato, T., Kajiwara, K., Hoshino, H., & Nakamura, T. (1997b). Quantitative Assessment of "Light Dark", "Deep-Pale" and "Heavy-Light" Colour Images. *Sen-I Gakkaishi*, *53*(1), 7–14.

Sato, T., Kajiwara, K., Hoshino, H., & Nakamura, T. (2000). Quantitative Evaluation and Categorising of Human Emotion Induced by Colour. *Advances in colour science and technology*, *3*(3), 1–7.

Schietecat, A. C., Lakens, D., Kort, Y. A. W. D., & Ijsselsteijn, W. A. (2014). Cross Modal Associations between Aggression and Light. *Proceedings Experiencing light 2014: international conference on the effects of light on wellbeing*, 1–4.

Schifferstein, H. (2006). The perceived importance of sensory modalities in product usage: A study of self-reports. *Acta Psychologica*, *121*(1), 41–64.

Schifferstein, H., & Cleiren, M. P. (2005). Capturing Product Experience - A split modality approach. *Acta Psychologica*, *118*(3), 293–318.

Schifferstein, H., Fenko, A., Desmet, P. M. A., Labbe, D., & Martin, N. (2013). Influence of package design on the dynamics of multisensory and emotional food experience. *Food Quality and Preference*, *27*(1), 18–25.

Schifferstein, H., Otten, J., Thoolen, F., & Hekkert, P. (2010). The experimental assessment of sensory dominance in a product development context. *Journal of Design Research*, *8*(2), 119–144.

Schifferstein, H., & Tanudjaja, I. (2004). Visualizing Fragrances Through Colours: The Mediating Role of Emotions. *Perception*, *33*(10), 1249–1266.

Seifert, C. E., Speechley, K. N., Booth, J. C., Stitt, L., & Gibson, M. C. (1996). A new analogue scale for assessing children' s pain : an initial validation study. *Children*, *64*(3), 435–443.

Semin, G. R., & Palma, T. A. (2014). Why the bride wears white: Grounding gender with brightness. *Journal of Consumer Psychology*, *24*(2), 217–225.

Sethi, I. K., Coman, I. L., & Stan, D. (2001). Mining association rules between low-level image features and high-level concepts. In *Proceeding of data mining and knowledge discovery: Theory, tools, and technology iii* (Vol. 4384, pp. 279–290).

Shankar, M. U., Levitan, C. A., Prescott, J., & Spence, C. (2009). The influence of color and label information on flavor perception. *Chemosensory Perception*, *2*(2), 53–58.

Shankar, M. U., Levitan, C. A., & Spence, C. (2010). Grape expectations: The role of cognitive influences in color-flavor interactions. *Consciousness and Cognition*, *19*(1), 380–390.

Shapiro, L. (2009). *Embodied Cognition*. Routledge.

Shepard, R. (1984). Ecological constraints on internal representation. *Psychological Review*, *91*(4), 417–447.

Sherman, G., & Clore, G. (2009). The Color of Sin: White and Black Are Perceptual Symbols of Moral Purity and Pollution. *Psychological science*, *20*(8), 1019–1025.

Slepian, M. L., & Ambady, N. (2014). Simulating sensorimotor metaphors: Novel metaphors influence embodied cognition. *Cognition*, *130*(3), 309–314.

Slepian, M. L., Masicampo, E. J., Toosi, N. R., & Ambady, N. (2012). The physical burdens of secrecy. *Journal of Experimental Psychology: General*, *141*(4), 619–624.

Smith–McLallen, A., Johnson, B. T., Dovidio, J. F., & Pearson, A. R. (2006). Black and White: The Role of Color Bias in Implicit Race Bias. *Social Cognition*, *24*(1), 46–73.

Snowden, R. J., Wichter, J., & Gray, N. S. (2008). Implicit and explicit measurements of sexual preference in gay and heterosexual men: A comparison of priming techniques and the Implicit Association Task. *Archives of Sexual Behavior*, *37*(4), 558–565.

Song, H., Vonasch, A. J., Meier, B. P., & Bargh, J. A. (2012). Brighten up: Smiles facilitate perceptual judgment of facial lightness. *Journal of Experimental Social Psychology*, *48*(1), 450–452.

Spence, C. (2002). The ICI report on the secret of the senses.

Spence, C. (2010). The multisensory perception of flavour. *Psychologist*, *23*(9), 720–723.

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, *73*(4), 971–995.

Spence, C., & Wan, X. (2015). Beverage perception and consumption: The influence of the container on the perception of the contents. *Food Quality and Preference*, *39*, 131–140.

Spence, C., Wan, X., Woods, A., Velasco, C., Deng, J., Youssef, J., & Deroy, O. (2015). On tasty colours and colourful tastes? Assessing, explaining, and utilizing crossmodal correspondences between colours and basic tastes. *Flavour*, *4*(23).

Stefanowitsch, A. (2006). Words and their metaphors: A corpus-based approach. *Trends in Linguistics Studies and Monographs*, *171*(63).

Stein, B. E., & Meredith, M. E. (1993). *The Merging of the Sense.* Mit Press Cambridge, MA.

Stone, L. S., & Thompson, P. (1992). Human speed perception is contrast dependent. *Vision Research*, *32*(8), 1535–1549.

Stone, N. J., & English, A. J. (1998). Task Type, Posters, and Workspace Color on Mood, Satisfaction, and Performance. *Journal of Environmental Psychology*, *18*(2), 175–185.

Sun, Y. (2012). Warm It Up with Love : The Effect. *Journal of Consumer Research*, *39*(2), 293–306.

Sunaga, T., Park, J., & Spence, C. (2016). Effects of Lightness-Location Congruency on Consumers' Purchase Decision-Making. *Psychology and Marketing*, *33*(11), 934–950.

Sutton, T. M., & Altarriba, J. (2015). Color associations to emotion and emotion-laden words: A collection of norms for stimulus construction and selection. *Behavior research methods*, *48*(2), 686–728.

Suzuki, M., & Gyoba, J. (2002). Contrastive analysis of sensory-relevance of factors affecting aesthetic impressions. *Tohoku psychologica folia*, *60*, 81–92.

Taft, C. (1997). Color Meaning and Context: Comparisons of Semantic Ratings of Colors on Samples and Objects. *Color Research & Application*, *22*(1), 40–50.

Takeuchi, T., De Valois, K. K., & Hardy, J. L. (2003). The influence of color on the perception of luminance motion. *Vision Research*, *43*(10), 1159–1175.

Talmy, L. (2005). The fundamental system of spatial schemas in language. *From Perception to Meaning: Image Schemas in Cognitive Linguistics*, 199–234.

Tanaka, J., & Presnell, L. (1999). Color diagnosticity in object recognition. *Perception & psychophysics*, *61*(6), 1140–1153.

Taylor, C. (1930). Visual perception versus visual plus kinaesthetic perception in judging colored weights. *The Journal of General Psychology*, *4*(1-4), 229.

Taylor, I. L., & Sumner, F. C. (1945). Actual Brightness and Distance of Individual Colors When Their Apparent Distance is Held Constant. *Journal of Psychology*, *19*(1), 79–85.

Tedford Jr, W. H., Bergquist, S. L., & Flynn, W. E. (1977). The size-color illusion. *The Journal of General Psychology*, *97*(1), 145–149.

Tharangie, K. G. D., Irfan, C. M. a., Marasinghe, C. a., & Yamada, K. (2008). Kansei engineering assessing system to enhance the usability in e-learning web interfaces: colour basis. In *Proceedings of the international conference on computers in educations* (Vol. 1, pp. 145–150).

Thomas, S., Joseph, C., & Laccetti, J. (2009). Transliteracy as a unifying perspective. *Handbook of Research on Social Software and Developing Community Ontologies*, 448–465.

Thorstenson, C. A., Pazda, A. D., & Elliot, A. J. (2015). Sadness Impairs Color Perception. *Psychological Science*, *26*(11), 1822.

Tinker, M. a. (1938). Effect of stimulus texture upon apparent warmth and affectuve value of colors. *The American journal of psychology*, *11*(3), 532–535.

Tobin, D. J. (2008). Human hair pigmentation-biological aspects. *International journal of cosmetic science*, *30*(4), 233–257.

Tolaas, J. (1991). Notes on the Origin of Some Spatialization Metaphors. *Metaphor and Symbolic Activity*, *6*(3), 203–218.

Valdez, P., & Mehrabian, A. (1994). Effects of color on emotions. *Journal of experimental psychology. General*, *123*(4), 394–409.

Van Essen, D. C., Felleman, D. J., DeYoe, E. A., Olavarria, J., & Knierim, J. (1990). Modular and hierarchical organization of extrastriate visual cortex in the macaque monkey. In *Cold spring harbor symposium on quantitative biology* (Vol. 55, pp. 679–696).

Van Ittersum, K., & Wansink, B. (2012). Plate Size and Color Suggestibility: The Delboeuf Illusion's Bias on Serving and Eating Behavior. *Journal of Consumer Research*, *39*(2), 215–228.

van Elk, M., van Schie, H. T., & Bekkering, H. (2010). From left to right: Processing acronyms referring to names of political parties activates spatial associations. *Quarterly journal of experimental psychology*, *63*(11), 2202–2019.

van Rompay, T. (2005). *Expressions: Embodiment in the experience of design.* Delft University of Technology.

Varela, F. J., Thompson, E., & Rosch, E. (1992). *The Embodied Mind: Cognitive Science and Human Experience*. MIT press.
Verhoeff, F. H. (1941). Occlusion hypertropia. *Archives of Ophthalmology*, *25*(5), 780–795.
Vess, M. (2012). Warm Thoughts: Attachment Anxiety and Sensitivity to Temperature Cues. *Psychological Science*, *23*(5), 472–474.
Vessel, E. A., & Rubin, N. (2010). Beauty and the beholder: highly individual taste for abstract, but not real-world images. *Journal of vision*, *10*(2), 1–14.
Viberg, Å. (1983). The verbs of perception: a typological study. *Linguistics*, *21*(1), 123–162.
Walker, L., Walker, P., & Francis, B. (2012). A common scheme for cross-sensory correspondences across stimulus domains. *Perception*, *41*(10), 1186–1192.
Walker, P., Francis, B., & Walker, L. (2010). The brightness-weight illusion: Darker objects look heavier but feel lighter. *Experimental Psychology*, *57*(6), 462–469.
Wallis, W. A. (1935). The influence of color on apparent size. *The Journal of General Psychology*, *13*(1), 193–199.
Wan, X., Woods, A. T., Salgado-Montejo, A., Velasco, C., & Spence, C. (2015). Assessing the expectations associated with pharmaceutical pill colour and shape. *Food Quality and Preference*, *45*, 171–182.
Wan, X., Woods, A. T., Van Den Bosch, J. J. F., McKenzie, K. J., Velasco, C., & Spence, C. (2014). Cross-cultural differences in crossmodal correspondences between basic tastes and visual features. *Frontiers in Psychology*, *5*.
Warden, C. J., & Flynn, E. L. (1931). The effect of color on apparent size and weight. *The American Journal of Psychology*, *43*(1), 109–111.
Warren, R., Gartstein, V., Kligman, A. M., Montagna, W., Allendorf, R. a., & Ridder, G. M. (1991). Age, sunlight, and facial skin: a histologic and quantitative study. *Journal of the American Academy of Dermatology*, *25*(5), 751–60.
Wastiels, L., Schifferstein, H. N. J., Heylighen, A., & Wouters, I. (2012a). Red or rough, what makes materials warmer? *Materials and Design*, *42*, 441–449.
Wastiels, L., Schifferstein, H. N. J., Heylighen, A., & Wouters, I. (2012b). Relating material experience to technical parameters: A case study on visual and tactile warmth perception of indoor wall materials. *Building and Environment*, *49*(1), 359–367.
Wastiels, L., Schifferstein, H. N. J., Wouters, I., & Heylighen, A. (2013). Touching materials visually: About the dominance of vision in building material assessment. *International Journal of Design*, *7*(2), 31–41.
Weller, L., & Livingston, R. (1988). Effect of Color of Questionnaire on Emotional Responses Effect of Color of Questionnaire on Emotional Responses. *The Journal of general psychology*, *115*(4), 433–440.
Wells, N. A. (1910). A description of the affective character of the colors of the spectrum. *Psychological Bulletin*, *7*(6), 181–195.
Wheatley. (1973). Putting colour into marketing. *Marketing*, *67*, 24–29.
Whitfield, T. W., & Whiltshire, T. J. (1990). Color psychology: A critical review. *Genetic, social, and general psychology monographs*.

Wierzbicka, A. (1990). The meaning of color terms: semantic, culture and cognition. *Cognitive Linguistics*, *1*(1), 99–150.

Wilkowski, B. M., Meier, B. P., Robinson, M. D., Carter, M. S., & Feltman, R. (2009). Hot-headed is more than an expression: the embodied representation of anger in terms of heat. *Emotion*, *9*(4), 464–477.

Williams, J., Boswell, D., & Best, D. (1975). Evaluative Responses of Preschool Children to the Colors White and Black. *Child Development*, *46*(2), 501–508.

Williams, L., & Bargh, J. A. (2008). Experiencing physical warmth promotes interpersonal warmth. *Science*, *322*(5901), 606–607.

Wilson, A., & Golonka, S. (2013). Embodied Cognition is Not What you Think it is. *Frontiers in Psychology*, *4*(58).

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic bulletin & review*, *9*(4), 625–636.

Wogalter, M. S., Conzola, V. C., & Smith-Jackson, T. L. (2002). Research based guidelines for warning design and evaluation. *Applied Economics*, *33*(3), 219–230.

Wong, D., & Baker, C. (1988). Pain in children: comparison of assessment scales. *Pediatric Nursing*, *14*(1), 9–17.

Woods, A. T., Spence, C., Butcher, N., & Deroy, O. (2013). Fast lemons and sour boulders: Testing crossmodal correspondences using an internet-based testing methodology. *i-Perception*, *4*(6), 365–379.

Wright, A. (2001). *The Beginner's Guide to Colour Psychology.* Kyle Cathie Limited.

Wright, B. (1962). The influence of hue, lightness, and saturation on apparent warmth and weight. *The American Journal of Psychology*, *75*(2), 232–241.

Xiao, K., Li, C. J., Luo, M. R., & Taylor, C. (2004). Colour Appearance for Dissimilar Sizes. In *Conference on colour in graphics, imaging, and vision* (pp. 12–16).

Xiao, K., Luo, M. R., Li, C., Cui, G., & Park, D. (2011). Investigation of colour size effect for colour appearance assessment. *Color Research and Application*, *36*(3), 201–209.

Xin, J. H., Cheng, K. M., Taylor, G., Sato, T., & Hansuebsai, A. (2004). Cross-regional comparison of colour emotions Part I: Quantitative analysis. *Color Research and Application*, *29*(6), 451–457.

Xu, A. J., & Labroo, A. A. (2014). Incandescent affect: Turning on the hot emotional system with bright light. *Journal of Consumer Psychology*, *24*(2), 207–216.

Yoo, H. S., & Smith-Jackson, T. L. (2011). Colour size illusion on liquid crystal displays and design guidelines for bioinformatics tools. *Behaviour & Information Technology*, *30*(6), 775–785.

Yoon, S.-y., & Wise, K. (2014). *Industrial Applications of Affective Engineering.* Springer.

Young, S. G., Elliot, A. J., Feltman, R., & Ambady, N. (2013). Red enhances the processing of facial expressions of anger. *Emotion*, *13*(3), 380–384.

Zachar, G., Schrott, A., & Kabai, P. (2008). Context-dependent prey avoidance in chicks persists following complete telencephalectomy. *Brain Research Bulletin*, *76*(3), 289–292.

Zampini, M., Sanabria, D., Phillips, N., & Spence, C. (2007). The multisensory perception of flavor: Assessing the influence of color cues on flavor discrimination responses. *Food Quality and Preference*, *18*(7), 975–984.

Zhong, C.-B., Bohns, V. K., & Gino, F. (2010). Good Lamps Are the Best Police: Darkness Increases Dishonesty and Self-Interested Behavior. *Psychological Science*, *21*(3), 311–314.

Ziat, M., Balcer, C. A., Shirtz, A., & Rolison, T. (2016). A century later, the hue-heat hypothesis: Does color truly affect temperature perception? *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*, 273–280.

A. Appendix

The appendix contains additional material used in the surveys and experiments and detailed results of statistical procedures as well as pre-tests. Asterisks indicating significance refer to corrected p-values (if applicable). All p-values are reported two-tailed, * $p < .05$, ** $p < .01$, *** $p < .001$. Note that reported p-values are rounded, but the number of asterisks refers to the not rounded p-value.

A.1. Study 1: Color-Image Schema Associations in German Subjects

A.1.1. Instructions (in German)

Auf den nachfolgenden Seiten sehen Sie verschiedene Farben. Ihre Aufgabe ist es, jede Farbe nach verschiedenen Eigenschafts-Dimensionen zu bewerten, z.B. warm-kalt, hell-dunkel. Können Sie eine Farbe bezüglich einer Dimension nicht einordnen, so wählen Sie bitte die mittlere Antwortoption aus.

Wichtig! Bitte stellen Sie jetzt Ihren Monitor auf maximale Helligkeit, um die vielen dunklen Farbtöne voneinander unterscheiden zu können.

A.1.2. Stimulus material

Table A.1.: *HSB Values, Related Color Patches and Color Names*

H	S	B	Color Patch	Color Name
212	35	100		Blue
212	100	100		Blue
212	90	25		Blue
212	35	35		Blue
0	35	100		Red
0	100	75		Red
0	100	25		Red
0	35	35		Red
56	35	100		Yellow
56	100	95		Yellow
56	100	35		Yellow
56	35	35		Yellow
36	35	100		Orange
36	100	100		Orange
36	100	35		Orange
36	35	35		Orange
282	35	100		Violet
282	100	75		Violet
282	100	25		Violet
282	35	35		Violet
130	35	90		Green
130	100	100		Green
130	75	25		Green
130	35	35		Green
35	40	50		Brown
35	80	50		Brown
35	80	25		Brown
35	40	25		Brown
181	35	100		Cyan
181	100	100		Cyan
181	100	35		Cyan
181	35	35		Cyan
n.d.	0	100		White
n.d.	0	50		Gray
n.d.	0	0		Black

A.1.3. Results

A.1.3.1. Correlation Coefficients

Table A.2.: *Pearson Correlation Coefficients r and p-Values for each Image Schema (Split-half Reliability)*

Image schema	$r(35)$	p
BIG-SMALL***	.879	.000
BRIGHT-DARK***	.992	.000
CLEAN-DIRTY***	.980	.000
FAR-NEAR***	.845	.000
FAST-SLOW***	.950	.000
FULL-EMPTY***	.936	.000
HARD-SOFT***	.956	.000
HEAVY-LIGHT***	.977	.000
LOUD-SILENT***	.932	.000
OLD-YOUNG***	.980	.000
PAINFUL-NOT PAINFUL***	.857	.000
SMELLS GOOD-SMELLS BAD***	.956	.000
SMOOTH-ROUGH***	.931	.000
STRONG-WEAK***	.941	.000
TASTES GOOD-TASTES BAD***	.881	.000
WARM-COLD***	.947	.000

A.1.3.2. Descriptive Data

Table A.3.: *Ms (SDs) for Image Schema Ratings of Different Colors*

	Image Schema				
HSB	BIG-SMALL	BRIGHT-DARK	CLEAN-DIRTY	FAR-NEAR	FAST-SLOW
212, 35, 100	4.15 (1.58)	2.17 (0.99)	4.45 (1.42)	4.45 (1.55)	2.89 (1.59)
212, 100, 100	3.48 (1.36)	3.22 (1.17)	4.97 (1.20)	4.97 (1.52)	2.61 (1.49)
212, 90, 25	2.80 (1.27)	5.93 (1.00)	4.85 (1.38)	4.85 (1.67)	3.82 (1.24)
212, 35, 35	3.27 (1.42)	5.30 (1.18)	4.49 (1.56)	4.49 (1.52)	3.70 (1.36)
0, 35, 100	4.93 (1.29)	2.39 (1.11)	5.71 (1.22)	5.71 (1.30)	3.00 (1.47)
0, 100, 75	2.96 (1.42)	3.84 (1.42)	6.29 (1.20)	6.29 (1.32)	1.89 (1.53)
0, 100, 25	2.91 (1.31)	5.77 (0.90)	5.69 (1.24)	5.69 (1.54)	3.71 (1.35)
0, 35, 35	3.62 (1.17)	5.17 (1.09)	5.18 (1.37)	5.18 (1.51)	3.96 (1.23)
56, 35, 100	4.05 (1.45)	1.94 (1.10)	5.26 (1.44)	5.26 (1.49)	2.98 (1.41)
56, 100, 95	3.53 (1.52)	2.09 (1.18)	5.82 (1.26)	5.82 (1.50)	1.90 (1.26)
56, 100, 35	3.60 (1.28)	5.20 (1.11)	4.80 (1.34)	4.80 (1.37)	3.79 (1.24)
56, 35, 35	3.70 (1.39)	5.12 (1.11)	4.50 (1.23)	4.50 (1.43)	4.05 (1.18)
36, 35, 100	4.35 (1.34)	2.12 (1.06)	5.44 (1.40)	5.44 (1.57)	3.27 (1.41)
36, 100, 100	3.59 (1.28)	2.87 (1.31)	5.87 (1.25)	5.87 (1.30)	2.42 (1.52)
36, 100, 35	3.42 (1.29)	5.33 (0.94)	4.99 (1.22)	4.99 (1.43)	4.04 (1.17)
36, 35, 35	3.65 (1.37)	5.39 (1.23)	4.76 (1.49)	4.76 (1.46)	4.12 (1.17)
282, 35, 100	4.66 (1.27)	2.32 (1.12)	5.25 (1.40)	5.25 (1.41)	3.25 (1.43)
282, 100, 75	3.46 (1.34)	4.35 (1.37)	5.34 (1.30)	5.34 (1.31)	3.07 (1.52)
282, 100, 25	3.13 (1.27)	5.87 (1.14)	4.95 (1.27)	5.26 (1.45)	3.79 (1.12)
282, 35, 35	3.79 (1.34)	4.97 (1.28)	4.20 (1.22)	4.95 (1.40)	3.59 (1.29)
130, 35, 90	4.46 (1.28)	2.30 (1.21)	4.92 (1.38)	4.92 (1.46)	2.90 (1.46)
130, 100, 100	3.63 (1.44)	2.26 (1.11)	5.61 (1.29)	5.61 (1.48)	1.43 (1.23)
130, 75, 25	3.24 (1.28)	5.48 (1.07)	5.30 (1.44)	5.30 (1.45)	3.84 (1.26)
130, 35, 35	3.57 (1.25)	5.01 (1.15)	4.89 (1.35)	4.89 (1.33)	3.79 (1.20)
35, 40, 50	3.99 (1.24)	4.29 (1.43)	2.79 (1.26)	3.29 (1.36)	4.04 (1.09)
35, 80, 50	3.84 (1.28)	4.79 (1.17)	2.64 (1.28)	2.93 (1.36)	3.94 (1.31)
35, 80, 25	3.14 (1.15)	5.79 (0.99)	2.33 (1.28)	3.07 (1.47)	4.08 (1.30)
35, 40, 25	3.06 (1.18)	5.81 (0.99)	2.54 (1.35)	3.05 (1.53)	4.18 (1.20)
181, 35, 100	4.28 (1.42)	2.10 (1.09)	4.63 (1.22)	4.63 (1.52)	2.64 (1.55)
181, 100, 100	3.87 (1.35)	2.41 (1.21)	5.03 (1.20)	5.03 (1.52)	2.12 (1.30)
181, 100, 35	3.31 (1.36)	5.03 (1.20)	4.99 (1.38)	4.99 (1.53)	3.51 (1.38)
181, 35, 35	3.46 (1.27)	5.20 (1.12)	4.61 (1.34)	4.61 (1.47)	3.77 (1.14)
N.D., 0, 100	3.34 (1.72)	1.39 (1.07)	4.65 (1.06)	4.65 (1.82)	2.58 (1.70)
N.D., 0, 50	3.71 (1.49)	4.22 (1.55)	4.21 (1.56)	4.21 (1.56)	4.07 (1.38)
N.D., 0, 0	2.51 (1.46)	6.84 (0.58)	5.14 (1.90)	5.14 (2.00)	3.26 (1.72)

Table A.3.: *Ms (SDs) for Image Schema Ratings of Different Colors (Continued)*

HSB		FULL-EMPTY	HARD-SOFT	HEAVY-LIGHT	LOUD-SILENT	OLD-YOUNG
212, 35, 100		3.88 (1.30)	3.84 (1.55)	5.63 (1.25)	4.83 (1.45)	5.12 (1.46)
212, 100, 100		2.39 (1.46)	2.90 (1.49)	4.58 (1.49)	3.58 (1.45)	5.12 (1.18)
212, 90, 25		1.88 (1.43)	3.10 (1.36)	2.51 (1.09)	4.01 (1.42)	3.55 (1.27)
212, 35, 35		2.54 (1.68)	2.93 (1.54)	2.89 (1.39)	4.11 (1.37)	3.54 (1.42)
0, 35, 100		2.92 (1.27)	4.36 (1.29)	5.41 (1.22)	4.58 (1.51)	5.20 (1.48)
0, 100, 75		1.46 (1.31)	2.37 (1.60)	3.30 (1.42)	2.28 (1.20)	5.10 (1.14)
0, 100, 25		1.38 (1.16)	3.15 (1.59)	2.47 (1.11)	3.44 (1.32)	3.38 (1.23)
0, 35, 35		2.27 (1.30)	3.05 (1.48)	3.12 (1.43)	4.36 (1.38)	3.02 (1.19)
56, 35, 100		3.55 (1.45)	4.01 (1.43)	5.46 (1.26)	4.63 (1.42)	4.88 (1.64)
56, 100, 95		1.91 (1.46)	2.97 (1.70)	4.88 (1.43)	2.75 (1.48)	5.36 (1.34)
56, 100, 35		2.29 (1.30)	2.69 (1.45)	2.96 (1.13)	4.06 (1.29)	2.89 (1.39)
56, 35, 35		2.82 (1.43)	2.64 (1.42)	2.91 (1.23)	4.33 (1.43)	2.68 (1.22)
36, 35, 100		3.51 (1.53)	4.31 1.25)	5.31 (1.11)	5.04 (1.29)	4.54 (1.53)
36, 100, 100		1.71 (1.26)	3.35 (1.46)	4.16 (1.38)	3.01 (1.33)	5.19 (1.34)
36, 100, 35		2.28 (1.35)	2.81 (1.43)	2.68 (1.24)	4.04 (1.27)	2.71 (1.32)
36, 35, 35		2.43 (1.50)	2.57 (1.50)	2.85 (1.26)	4.25 (1.46)	2.43 (1.16)
282, 35, 100		3.21 (1.35)	4.11 (1.46)	5.25 (1.20)	4.72 (1.46)	5.09 (1.54)
282, 100, 75		2.04 (1.31)	2.98 (1.41)	3.43 (1.42)	3.24 (1.39)	4.74 (1.34)
282, 100, 25		1.66 (1.23)	2.77 (1.54)	2.72 (1.22)	4.04 (1.39)	3.65 (1.28)
282, 35, 35		2.39 (1.36)	3.15 (1.36)	3.15 (1.37)	5.04 (1.46)	3.82 (1.36)
130, 35, 90		3.42 (1.37)	3.49 (1.54)	5.41 (1.21)	4.62 (1.71)	5.02 (1.36)
130, 100, 100		2.22 (1.37)	2.42 (1.51)	5.03 (1.37)	2.59 (1.43)	5.72 (1.20)
130, 75, 25		1.87 (1.28)	3.08 (1.44)	2.97 (1.27)	3.82 (1.38)	3.45 (1.33)
130, 35, 35		2.62 (1.46)	3.01 (1.36)	3.28 (1.31)	4.40 (1.23)	3.33 (1.32)
35, 40, 50		3.08 (1.41)	3.12 (1.54)	3.71 (1.41)	3.29 (1.28)	3.14 (1.31)
35, 80, 50		2.23 (1.22)	2.94 (1.33)	3.13 (1.28)	3.73 (1.34)	2.75 (1.31)
35, 80, 25		1.94 (1.30)	2.46 (1.49)	2.54 (1.08)	4.02 (1.42)	2.50 (1.09)
35, 40, 25		2.19 (1.53)	2.42 (1.54)	2.71 (1.32)	3.92 (1.43)	2.66 (1.33)
181, 35, 100		3.68 (1.45)	3.38 (1.52)	5.52 (1.20)	4.37 (1.57)	5.35 (1.29)
181, 100, 100		2.97 (1.46)	2.99 (1.59)	5.26 (1.35)	3.42 (1.52)	5.50 (1.18)
181, 100, 35		1.98 (1.32)	3.17 (1.50)	3.21 (1.39)	3.86 (1.41)	4.23 (1.44)
181, 35, 35		2.67 (1.50)	2.91 (1.39)	3.18 (1.38)	4.55 (1.15)	3.46 (1.40)
N.D., 0, 100		4.26 (1.89)	3.53 (2.00)	5.98 (1.33)	5.04 (1.66)	4.88 (1.62)
N.D., 0, 50		3.63 (1.62)	2.51 (1.62)	3.89 (1.74)	4.87 (1.47)	2.80 (1.40)
N.D., 0, 0		2.08 (2.17)	1.47 (1.49)	1.94 (1.27)	3.46 (1.89)	3.57 (1.40)

Table A.3.: *Ms (SDs) for Image Schema Ratings of Different Colors (Continued)*

		Image Schema			
		PAINFUL-	SMELLS GOOD-	SMOOTH-	STRONG-
HSB		NOT PAINFUL	SMELLS BAD	ROUGH	WEAK
212, 35, 100		5.13 (1.43)	3.07 (1.41)	1.78 (1.35)	3.79 (1.44)
212, 100, 100		4.79 (1.45)	3.27 (1.19)	1.92 (1.14)	2.08 (1.32)
212, 90, 25		4.69 (1.41)	3.73 (1.18)	2.80 (1.43)	1.54 (1.08)
212, 35, 35		4.39 (1.60)	4.08 (1.20)	2.99 (1.61)	2.37 (1.49)
0, 35, 100		4.92 (1.67)	3.07 (1.52)	1.73 (1.15)	3.69 (1.48)
0, 100, 75		3.26 (1.65)	3.63 (1.40)	2.02 (1.40)	1.04 (1.29)
0, 100, 25		4.43 (1.69)	3.56 (1.30)	2.72 (1.54)	1.42 (1.14)
0, 35, 35		4.50 (1.43)	4.12 (1.39)	3.29 (1.39)	2.71 (1.47)
56, 35, 100		5.08 (1.40)	3.21 (1.45)	2.06 (1.43)	3.66 (1.53)
56, 100, 95		4.26 (1.63)	3.13 (1.39)	1.86 (1.34)	1.56 (1.27)
56, 100, 35		3.98 (1.31)	5.22 (1.33)	3.21 (1.43)	2.84 (1.44)
56, 35, 35		4.05 (1.46)	5.28 (1.14)	3.94 (1.32)	2.96 (1.46)
36, 35, 100		5.32 (1.27)	3.08 (1.40)	2.05 (1.43)	3.83 (1.29)
36, 100, 100		4.54 (1.56)	3.03 (1.42)	2.22 (1.37)	1.57 (1.25)
36, 100, 35		4.09 (1.47)	5.10 (1.38)	3.47 (1.39)	2.44 (1.31)
36, 35, 35		3.97 (1.49)	5.07 (1.35)	3.77 (1.51)	2.68 (1.53)
282, 35, 100		5.00 (1.48)	3.17 (1.38)	1.79 (1.13)	4.00 (1.46)
282, 100, 75		4.38 (1.45)	3.52 (1.32)	2.42 (1.33)	2.14 (1.54)
282, 100, 25		4.56 (1.49)	3.52 (1.22)	2.60 (1.40)	1.58 (1.08)
282, 35, 35		4.55 (1.34)	3.77 (1.41)	2.98 (1.41)	2.67 (1.30)
130, 35, 90		4.74 (1.51)	3.56 (1.55)	2.14 (1.41)	3.59 (1.43)
130, 100, 100		3.59 (1.71)	4.06 (1.54)	1.82 (1.27)	1.71 (1.44)
130, 75, 25		4.75 (1.41)	3.57 (1.42)	2.89 (1.45)	1.93 (1.19)
130, 35, 35		4.61 (1.40)	4.21 (1.38)	3.36 (1.31)	2.78 1.32)
35, 40, 50		4.50 (1.49)	4.99 (1.23)	3.41 (1.41)	3.54 (1.38)
35, 80, 50		4.45 (1.24)	4.95 (1.44)	3.45 (1.39)	2.82 (1.34)
35, 80, 25		4.11 (1.52)	5.27 (1.48)	3.65 (1.37)	2.03 (1.21)
35, 40, 25		3.96 (1.62)	5.02 (1.46)	3.94 (1.35)	2.57 (1.46)
181, 35, 100		4.63 (1.51)	3.49 (1.48)	1.86 (1.35)	3.29 (1.54)
181, 100, 100		4.49 (1.62)	3.64 (1.34)	1.58 (1.20)	2.54 (1.49)
181, 100, 35		4.96 (1.40)	3.67 (1.35)	2.71 (1.46)	2.07 (1.27)
181, 35, 35		4.54 (1.52)	4.16 (1.27)	3.22 (1.52)	2.96 (1.33)
N.D., 0, 100		4.83 (1.65)	3.16 (1.37)	1.32 (1.32)	2.90 (1.91)
N.D., 0, 50		4.41 (1.53)	4.96 (1.20)	3.78 (1.77)	3.57 (1.67)
N.D., 0, 0		3.71 (1.78)	4.70 (1.47)	2.12 (1.90)	0.69 (1.07)

Table A.3.: *Ms (SDs) for Image Schema Ratings of Different Colors (Continued)*

		Image Schema	
		TASTES GOOD-	WARM-
HSB		TASTES BAD	COLD
212, 35, 100		3.93 (1.41)	4.20 (1.24)
212, 100, 100		3.96 (1.35)	3.99 (1.33)
212, 90, 25		4.27 (1.24)	3.97 (1.39)
212, 35, 35		4.34 (1.18)	4.01 (1.37)
0, 35, 100		3.78 (1.41)	2.01 (1.23)
0, 100, 75		3.73 (1.51)	1.48 (1.36)
0, 100, 25		3.57 (1.40)	1.68 (1.33)
0, 35, 35		4.36 (1.39)	2.61 (1.35)
56, 35, 100		3.86 (1.38)	2.48 (1.45)
56, 100, 95		3.63 (1.40)	1.78 (1.39)
56, 100, 35		5.45 (1.37)	3.21 (1.32)
56, 35, 35		5.36 (1.22)	3.46 (1.46)
36, 35, 100		3.65 (1.41)	2.41 (1.51)
36, 100, 100		3.45 (1.45)	1.29 (1.19)
36, 100, 35		5.20 (1.42)	2.68 (1.31)
36, 35, 35		5.24 (1.33)	3.27 (1.42)
282, 35, 100		4.05 (1.52)	2.73 (1.32)
282, 100, 75		4.42 (1.43)	3.10 (1.34)
282, 100, 25		4.14 (1.33)	3.04 (1.41)
282, 35, 35		4.08 (1.30)	2.94 (1.45)
130, 35, 90		4.13 (1.60)	3.55 (1.41)
130, 100, 100		4.40 (1.64)	3.41 (1.38)
130, 75, 25		3.96 (1.28)	2.89 (1.39)
130, 35, 35		4.54 (1.31)	3.35 1.39)
35, 40, 50		5.11 (1.18)	3.03 (1.36)
35, 80, 50		5.00 (1.32)	2.48 (1.28)
35, 80, 25		5.06 (1.57)	2.94 (1.58)
35, 40, 25		5.25 (1.31)	3.12 (1.50)
181, 35, 100		4.18 (1.31)	4.20 (1.29)
181, 100, 100		4.10 (1.57)	4.14 (1.34)
181, 100, 35		4.09 (1.41)	3.45 (1.47)
181, 35, 35		4.59 (1.33)	3.71 (1.39)
N.D., 0, 100		3.78 (1.32)	3.84 (1.65)
N.D., 0, 50		5.29 (1.16)	4.24 (1.30)
N.D., 0, 0		5.02 (1.34)	4.06 (1.60)

A.1.3.3. Regression Models

Table A.4.: *Color Regression Models of 16 Bipolar Image Schema Ratings (German Subjects)*

Image Schema	Intercept (Red)	Hue	Saturation (S)	Brightness (B)
BIG-SMALL***	0.067	+ 0.047 (Yellow) - 0.175 (Blue) + 0.045 (Cyan) + 0.035 (Green) + 0.071 (Orange) + 0.130 (Violet) - 0.809 (Black)*** -0.302 (Gray) -0.933 (White)***	- 0.008*S***	+ 0.008* B***
BRIGHT-DARK***	1.867	- 0.373 (Yellow) + 0.018 (Blue) - 0.267 (Cyan) - 0.733 (Green)* - 0.126 (Orange) - 0.024 (Violet) + 0.528 (Black) + 0.011 (Gray) - 0.658 (White)	+ 0.006* S	- 0.028* B***
CLEAN-DIRTY***	1.336	+ 0.499 (Yellow) - 0.191 (Blue) - 0.113 (Cyan) + 0.049 (Green) + 0.557 (Orange)* - 0.039 (Violet) - 0.860 (Black) + 0.473 (Gray) - 0.860 (White)	- 0.002*S	- 0.020*B***
FAR-NEAR***	- 0.119	- 0.437 (Yellow)** - 0.692 (Blue)*** - 0.629 (Cyan)*** - 0.337 (Green)** - 0.323 (Orange)* - 0.338 (Violet)** + 0.151 (Black) - 0.661 (Gray)** - 0.553 (White)**	+ 0.005*S***	+ 0.004*B**

Table A.4.: *Color Regression Models of 16 Bipolar Image Schema Ratings (German Subjects) (Continued)*

Image Schema	Intercept (Red)	Hue	Saturation (S)	Brightness (B)
FAST-SLOW***	0.986	+ 0.098 (Yellow) + 0.119 (Blue) - 0.009 (Cyan) - 0.131 (Green) + 0.302 (Orange) + 0.172 (Violet) - 1.185 (Black)** + 0.089 (Gray) - 0.303 (White)	- 0.006*S**	- 0.014*B***
FULL-EMPTY***	- 0.708	+ 0.378 (Yellow)* + 0.243 (Blue) + 0.359 (Cyan)* + 0.121 (Green) + 0.271 (Orange) + 0.073 (Violet) - 0.346 (Black) - 0.558 (Gray)* - 0.558 (White)	- 0.011*S***	+ 0.009*B***
HARD-SOFT**	- 0.195	- 0.142 (Yellow) - 0.065 (Blue) - 0.124 (Cyan) - 0.211 (Green) - 0.010 (Orange) + 0.029 (Violet) - 1.300 (Black)** - 0.841 (Gray)* - 0.397 (White)	- 0.006*S**	+ 0.006*B**
HEAVY-LIGHT***	- 1.049	+ 0.227 (Yellow) + 0.134 (Blue) + 0.371 (Cyan)* + 0.368 (Green)* + 0.008 (Orange) + 0.105 (Violet) - 0.106 (Black) + 0.120 (Gray) + 0.447 (White)	- 0.006*S***	+ 0.024*B***

Table A.4.: *Color Regression Models of 16 Bipolar Image Schema Ratings (German Subjects) (Continued)*

Image Schema	Intercept (Red)	Hue	Saturation (S)	Brightness (B)
LOUD-SILENT**	0.711	+ 0.205 (Yellow) + 0.306 (Blue) + 0.283 (Cyan) + 0.061 (Green) + 0.306 (Orange) + 0.179 (Violet) - 1.093 (Black)* - 0.007 (Gray) + 0.202 (White)	- 0.012*S***	- 0.002*B
OLD-YOUNG***	- 1.121	- 0.266 (Yellow) + 0.046 (Blue) + 0.202 (Cyan) + 0.150 (Green) - 0.437 (Orange)** + 0.161 (Violet) + 1.147 (Black)*** - 0.463 (Gray) - 0.008 (White)	+ 0.004*S**	+ 0.021*B***
PAINFUL-NOT PAINFUL	0.675	+ 0.036 (Yellow) + 0.324 (Blue) + 0.260 (Cyan) + 0.077 (Green) + 0.136 (Orange) + 0.256 (Violet) - 0.547 (Black) - 0.151 (Gray) + 0.041 (White)	- 0.005*S	+ 0.004*B
SMELLS GOOD-SMELLS BAD*	0.711	+ 0.495 (Yellow) - 0.007 (Blue) + 0.167 (Cyan) + 0.191 (Green) + 0.394 (Orange) - 0.097 (Violet) + 0.134 (Black) + 0.817 (Gray) + 0.007 (White)	- 0.001*S	- 0.010*B***

Table A.4.: *Color Regression Models of 16 Bipolar Image Schema Ratings (German Subjects) (Continued)*

Image Schema	Intercept (Red)	Hue	Saturation (S)	Brightness (B)
SMOOTH-ROUGH***	0.272	+ 0.306 (Yellow)* - 0.005 (Blue) + 0.016 (Cyan) + 0.078 (Green) + 0.386 (Orange)** - 0.029 (Violet) - 1.296 (Black)*** + 0.596 (Gray)* - 0.497 (White)*	- 0.003*S**	- 0.014*B***
STRONG-WEAK***	0.438	+ 0.360 (Yellow) + 0.105 (Blue) + 0.323 (Cyan) + 0.105 (Green) + 0.154 (Orange) + 0.292 (Violet) - 1.759 (Black)*** + 0.017 (Gray) - 0.776 (White)	- 0.015*S***	+ 0.006*B**
TASTES GOOD-TASTES BAD	0.740	+ 0.550 (Yellow)* + 0.212 (Blue) + 0.316 (Cyan) + 0.286 (Green) + 0.413 (Orange) + 0.210 (Violet) + 0.337 (Black) + 0.856 (Gray)* + 0.085 (White)	- 0.002*S	- 0.006*B**
WARM-COLD***	- 0.728	+ 0.584 (Yellow)* + 1.520 (Blue)*** + 1.415 (Cyan)*** + 0.961 (Green)** + 0.352 (Orange) + 0.723 (Violet)** + 1.109 (Black)* + 1.370 (Gray)** + 1.210 (White)**	- 0.004*S	- 0.003*B

A.2. Study 2: Color-Image Schema Associations in Japanese Subjects

A.2.1. Instructions (in Japanese)

次ページ以降ではさまざまな色が表示されます．それぞれの画面で表示される色が「熱-冷」や「明-暗」などのことばの対のどちらが相応しいかについて評価して下さい．どちらも相応しくない場合には，中間を選んでください． 重要! 暗い色の違いが区別できるように，モニタの輝度を最大に設定してください

A.2.2. Stimulus material

Table A.5.: *Image schemas Grouped by Similarity and their Japanese Translation*

Group	Image Schema	Japanese Translation
Space	FAR-NEAR	遠-近
Containment	FULL-EMPTY	満-空
Attribute	BIG-SMALL	大-小
	BRIGHT-DARK	明-暗
	CLEAN-DIRTY	きれい-汚い
	FAST-SLOW	速-遅
	HARD-SOFT	硬-柔
	HEAVY-LIGHT	重-軽
	LOUD-SILENT	うるさい-静かな
	OLD-YOUNG	老-若
	PAINFUL-NOT PAINFUL	痛い-痛くない
	SMELLS GOOD-SMELLS BAD	香しい（いいにおいがする)-臭い
	SMOOTH-ROUGH	細-粗
	STRONG-WEAK	強-弱
	TASTES GOOD-TASTES BAD	美味しい-不味い
	WARM-COLD	暖-寒

A.2.2.1. Descriptive Data

Table A.6.: *Ms (SDs) for Image Schema Ratings of Different Colors*

		Image Schema				
HSB		BIG-SMALL	BRIGHT-DARK	CLEAN-DIRTY	FAR-NEAR	FAST-SLOW
212, 35, 100		3.86 (1.74)	2.38 (1.45)	2.42 (1.43)	4.34 (1.60)	2.22 (1.56)
212, 100, 100		3.63 (1.67)	2.52 (1.32)	2.38 (1.29)	4.89 (1.69)	1.74 (1.50)
212, 90, 25		2.79 (1.30)	5.96 (1.00)	4.18 (1.45)	4.40 (1.70)	4.14 (1.41)
212, 35, 35		3.48 (1.51)	5.82 (1.05)	4.93 (1.31)	4.14 (1.63)	4.08 (1.45)
0, 35, 100		4.51 (1.74)	2.04 (1.14)	2.29 (1.12)	5.74 (1.51)	2.90 (1.61)
0, 100, 75		2.77 (1.57)	2.59 (1.59)	2.86 (1.41)	5.92 (1.58)	2.04 (1.66)
0, 100, 25		2.84 (1.36)	5.33 (1.49)	4.59 (1.80)	4.56 (1.77)	3.99 (1.56)
0, 35, 35		3.49 (1.41)	5.26 (1.32)	5.07 (1.29)	4.81 (1.44)	4.30 (1.13)
56, 35, 100		4.49 (1.74)	1.89 (1.06)	2.45 (1.27)	5.36 (1.70)	2.70 (1.69)
56, 100, 95		4.11 (1.68)	2.12 (1.60)	2.77 (1.50)	6.10 (1.44)	2.00 (1.57)
56, 100, 35		3.38 (1.53)	5.38 (1.28)	5.67 (1.29)	4.63 (1.49)	4.22 (1.23)
56, 35, 35		3.30 (1.45)	5.68 (1.22)	5.62 (1.34)	4.36 (1.42)	4.29 (1.28)
36, 35, 100		4.67 (1.40)	2.40 (1.19)	2.99 (1.14)	5.37 (1.55)	3.12 (1.56)
36, 100, 100		3.16 (1.25)	2.21 (1.28)	2.71 (1.34)	5.36 (1.41)	2.77 (1.58)
36, 100, 35		3.03 (1.45)	5.47 (1.23)	5.71 (1.28)	4.97 (1.65)	4.41 (1.32)
36, 35, 35		3.32 (1.50)	5.66 (1.05)	5.30 (1.34)	4.21 (1.59)	4.37 (1.32)
282, 35, 100		4.59 (1.50)	2.82 (1.51)	2.96 (1.50)	4.81 (1.60)	3.04 (1.53)
282, 100, 75		3.18 (1.25)	4.52 (1.54)	3.71 (1.39)	5.32 (1.38)	3.44 (1.65)
282, 100, 25		3.49 (1.64)	5.73 (1.32)	3.93 (1.62)	4.75 (1.73)	4.10 (1.55)
282, 35, 35		3.47 (1.32)	5.51 (1.24)	4.53 (1.42)	4.48 (1.45)	4.01 (1.34)
130, 35, 90		4.16 (1.53)	2.41 (1.25)	2.52 (1.51)	5.16 (1.67)	2.38 (1.54)
130, 100, 100		4.01 (1.65)	1.90 (1.20)	2.33 (1.20)	5.88 (1.61)	2.04 (1.42)
130, 75, 25		3.42 (1.52)	5.41 (1.31)	4.47 (1.66)	5.07 (1.50)	4.04 (1.46)
130, 35, 35		3.63 (1.45)	5.32 (1.27)	4.85 (1.48)	4.25 (1.49)	4.04 (1.23)
35, 40, 50		3.74 (1.39)	5.04 (1.29)	2.85 (1.29)	3.42 (1.45)	4.01 (1.21)
35, 80, 50		3.47 (1.52)	5.07 (1.50)	2.77 (1.53)	2.93 (1.61)	4.71 (1.24)
35, 80, 25		2.79 (1.41)	5.92 (1.23)	2.38 (1.16)	3.04 (1.82)	4.73 (1.18)
35, 40, 25		2.90 (1.33)	6.00 (1.26)	2.23 (1.12)	3.22 (1.68)	4.44 (1.36)
181, 35, 100		4.55 (1.90)	2.07 (1.29)	2.16 (1.39)	4.96 (1.75)	1.97 (1.61)
181, 100, 100		3.85 (1.84)	1.89 (1.25)	2.03 (1.14)	5.60 (1.74)	1.70 (1.46)
181, 100, 35		3.45 (1.40)	5.12 (1.37)	4.04 (1.62)	4.44 (1.42)	3.81 (1.33)
181, 35, 35		3.63 (1.33)	5.22 (1.37)	4.66 (1.60)	4.21 (1.55)	3.93 (1.38)
N.D., 0, 100		3.26 (2.03)	1.55 (1.21)	1.84 (1.35)	5.01 (2.17)	1.75 (1.64)
N.D., 0, 50		4.03 (1.69)	4.86 (1.71)	4.75 (1.42)	3.99 (1.54)	3.84 (1.63)
N.D., 0, 0		2.07 (1.45)	6.52 (1.34)	4.55 (1.85)	4.30 (2.06)	3.58 (2.14)

Table A.6.: *Ms (SDs) for Image Schema Ratings of Different Colors (Continued)*

		Image Schema				
		FULL-	HARD-	HEAVY-	LOUD-	OLD-
HSB		EMPTY	SOFT	LIGHT	SILENT	YOUNG
212, 35, 100		4.44 (1.58)	2.86 (1.54)	5.77 (1.30)	5.11 (1.58)	5.29 (1.44)
212, 100, 100		3.62 (1.90)	3.49 (1.50)	4.77 (1.56)	4.52 (1.78)	5.34 (1.30)
212, 90, 25		2.25 (1.67)	5.08 (1.42)	2.42 (1.33)	5.47 (1.36)	3.03 (1.36)
212, 35, 35		2.55 (1.65)	4.64 (1.48)	2.85 (1.49)	5.42 (1.25)	2.82 (1.19)
0, 35, 100		2.64 (1.62)	2.36 (1.50)	5.58 (1.36)	3.67 (1.64)	5.96 (1.14)
0, 100, 75		1.42 (1.43)	4.66 (1.53)	3.18 (1.54)	2.15 (1.41)	5.30 (1.50)
0, 100, 25		1.66 (1.19)	4.90 (1.65)	2.41 (1.48)	4.60 (1.67)	2.99 (1.54)
0, 35, 35		2.33 (1.56)	3.90 (1.71)	2.81 (1.42)	4.82 (1.44)	2.70 (1.37)
56, 35, 100		3.41 (1.79)	2.40 (1.28)	5.56 (1.40)	3.73 (1.70)	5.23 (1.64)
56, 100, 95		2.45 (1.68)	3.71 (1.80)	5.26 (1.43)	1.99 (1.37)	5.63 (1.36)
56, 100, 35		2.56 (1.44)	3.75 (1.63)	2.82 (1.33)	4.81 (1.42)	2.23 (1.17)
56, 35, 35		2.63 (1.49)	4.01 (1.64)	2.77 (1.54)	5.19 (1.34)	2.42 (1.29)
36, 35, 100		3.25 (1.51)	2.41 (1.39)	5.44 (1.36)	4.32 (1.58)	4.81 (1.49)
36, 100, 100		2.26 (1.58)	3.26 (1.37)	4.27 (1.49)	2.79 (1.48)	5.00 (1.56)
36, 100, 35		2.22 (1.33)	3.95 (1.77)	2.52 (1.19)	4.67 (1.58)	2.52 (1.39)
36, 35, 35		2.66 (1.55)	4.66 (1.53)	2.77 (1.44)	5.00 (1.33)	2.41 (1.24)
282, 35, 100		3.14 (1.59)	2.89 (1.41)	4.99 (1.48)	4.00 (1.81)	4.78 (1.60)
282, 100, 75		2.51 (1.61)	4.16 (1.39)	3.33 (1.46)	3.81 (1.8)	3.51 (1.54)
282, 100, 25		2.08 (1.37)	4.56 (1.68)	2.62 (1.50)	4.40 (1.69)	3.08 (1.53)
282, 35, 35		2.63 (1.53)	4.33 (1.55)	2.92 (1.58)	4.85 (1.42)	2.71 (1.13)
130, 35, 90		3.84 (1.65)	2.51 (1.30)	5.44 (1.52)	4.73 (1.69)	5.27 (1.69)
130, 100, 100		3.14 (1.60)	3.36 (1.57)	5.36 (1.27)	3.34 (1.75)	6.11 (1.03)
130, 75, 25		2.18 (1.33)	3.64 (1.53)	2.85 (1.33)	5.10 (1.23)	3.38 (1.62)
130, 35, 35		2.52 (1.43)	3.85 (1.50)	3.18 (1.61)	5.11 (1.35)	3.05 (1.48)
35, 40, 50		2.85 (1.45)	3.81 (1.60)	3.68 (1.62)	3.33 (1.46)	2.68 (1.42)
35, 80, 50		2.12 (1.49)	3.89 (1.75)	2.78 (1.47)	3.48 (1.51)	2.32 (1.36)
35, 80, 25		1.97 (1.60)	4.23 (1.79)	2.33 (1.34)	3.27 (1.46)	2.29 (1.23)
35, 40, 25		2.45 (1.71)	4.73 (1.81)	2.30 (1.18)	3.10 (1.50)	2.07 (0.96)
181, 35, 100		4.44 (1.40)	2.95 (1.42)	5.93 (1.31)	4.81 (1.62)	5.74 (1.09)
181, 100, 100		3.67 (1.82)	3.25 (1.64)	5.53 (1.39)	4.71 (1.89)	5.74 (1.28)
181, 100, 35		2.78 (1.50)	4.21 (1.49)	3.00 (1.41)	5,04 (1.42)	3.30 (1.32)
181, 35, 35		2.64 (1.31)	4.26 (1.59)	3.12 (1.39)	5.03 (1.39)	3.03 (1.5)
N.D., 0, 100		4.44 (2.14)	2.81 (1.93)	6.19 (1.25)	5.55 (1.77)	5.41 (1.46)
N.D., 0, 50		3.97 (1.65)	4.32 (1.93)	3.84 (1.84)	5.38 (1.43)	3.18 (1.52)
N.D., 0, 0		2.07 (2.20)	5.71 (1.59)	1.77 (1.32)	5.19 (2.09)	3.71 (1.56)

Table A.6.: *Ms (SDs) for Image Schema Ratings of Different Colors (Continued)*

HSB		Image Schema: PAINFUL-NOT PAINFUL	SMELLS GOOD-SMELLS BAD	SMOOTH-ROUGH	STRONG-WEAK
212, 35, 100		5.01 (1.82)	3.07 (1.27)	2.21 (1.66)	4.12 (1.51)
212, 100, 100		4.66 (1.61)	3.32 (1.01)	2.44 (1.47)	2.47 (1.76)
212, 90, 25		4.40 (1.61)	4.32 (1.21)	3.63 (1.63)	2.26 (1.5)
212, 35, 35		4.60 (1.51)	4.77 (1.12)	3.56 (1.52)	2.58 (1.56)
0, 35, 100		5.38 (1.52)	2.60 (1.28)	2.03 (1.39)	3.96 (1.61)
0, 100, 75		2.04 (1.26)	3.23 (1.27)	3.90 (1.61)	0.88 (1.35)
0, 100, 25		3.62 (1.8)	3.82 (1.47)	3.75 (1.66)	1.78 (1.46)
0, 35, 35		4.36 (1.62)	4.40 (1.56)	3.60 (1.53)	3.04 (1.58)
56, 35, 100		5.21 (1.47)	3.08 (1.45)	2.47 (1.55)	4.05 (1.49)
56, 100, 95		3.73 (1.80)	3.22 (1.38)	2.48 (1.67)	2.00 (1.88)
56, 100, 35		4.63 (1.53)	4.93 (1.52)	4.00 (1.34)	2.93 (1.53)
56, 35, 35		4.49 (1.47)	4.97 (1.50)	4.05 (1.38)	2.89 (1.50)
36, 35, 100		5.51 (1.27)	3.12 (1.25)	2.29 (1.42)	4.33 (1.25)
36, 100, 100		4.05 (1.55)	2.63 (1.25)	3.00 (1.52)	2.25 (1.62)
36, 100, 35		4.48 (1.61)	5.08 (1.72)	3.92 (1.59)	2.96 (1.75)
36, 35, 35		4.36 (1.68)	5.07 (1.35)	3.88 (1.50)	2.93 (1.63)
282, 35, 100		4.67 (1.70)	3.26 (1.49)	2.40 (1.37)	3.49 (1.53)
282, 100, 75		3.75 (1.50)	3.85 (1.28)	3.55 (1.37)	2.27 (1.53)
282, 100, 25		4.14 (1.42)	4.01 (1.35)	3.40 (1.71)	2.15 (1.53)
282, 35, 35		4.38 (1.49)	4.33 (1.39)	3.45 (1.55)	2.52 (1.54)
130, 35, 90		4.99 (1.67)	2.93 (1.20)	2.29 (1.61)	3.95 (1.45)
130, 100, 100		4.53 (1.75)	3.14 (1.40)	2.29 (1.61)	2.71 (1.80)
130, 75, 25		5.03 (1.44)	4.15 (1.52)	3.42 (1.60)	2.49 (1.57)
130, 35, 35		4.63 (1.48)	4.58 (1.45)	3.58 (1.41)	2.96 (1.55)
35, 40, 50		4.73 (1.56)	4.86 (1.37)	3.58 (1.41)	3.48 (1.45)
35, 80, 50		4.37 (1.51)	4.73 (1.59)	3.96 (1.55)	2.85 (1.53)
35, 80, 25		4.37 (1.55)	4.71 (1.50)	4.14 (1.44)	2.48 (1.64)
35, 40, 25		4.08 (1.60)	4.89 (1.58)	4.01 (1.55)	2.58 (1.73)
181, 35, 100		5.10 (1.62)	3.11 (1.28)	1.79 (1.47)	4.33 (1.20)
181, 100, 100		4.88 (1.64)	3.07 (1.25)	1.77 (1.43)	3.36 (1.70)
181, 100, 35		4.63 (1.39)	4.22 (1.32)	3.18 (1.47)	3.08 (1.38)
181, 35, 35		4.78 (1.48)	4.77 (1.18)	3.56 (1.52)	2.92 (1.47)
N.D., 0, 100		5.40 (1.62)	3.30 (1.30)	1.59 (1.73)	3.19 (2.14)
N.D., 0, 50		4.36 (1.67)	4.36 (1.21)	3.55 (1.78)	3.64 (1.81)
N.D., 0, 0		3.53 (1.95)	4.29 (1.40)	3.66 (1.98)	1.18 (1.75)

Table A.6.: *Ms (SDs) for Image Schema Ratings of Different Colors (Continued)*

		Image Schema	
HSB		TASTES GOOD-TASTES BAD	WARM-COLD
212, 35, 100		3.90 (1.55)	4.59 (1.22)
212, 100, 100		4.36 (1.59)	4.63 (1.29)
212, 90, 25		5.18 (1.32)	4.55 (1.28)
212, 35, 35		5.33 (1.15)	4.36 (1.14)
0, 35, 100		2.74 (1.47)	1.55 (1.45)
0, 100, 75		3.40 (1.42)	0.99 (1.19)
0, 100, 25		4.48 (1.78)	2.32 (1.52)
0, 35, 35		4.77 (1.62)	2.79 (1.51)
56, 35, 100		3.07 (1.41)	1.90 (1.30)
56, 100, 95		3.55 (1.46)	1.85 (1.23)
56, 100, 35		5.03 (1.69)	3.40 (1.35)
56, 35, 35		5.49 (1.29)	3.51 (1.26)
36, 35, 100		3.44 (1.50)	1.86 (1.31)
36, 100, 100		2.77 (1.40)	1.25 (1.17)
36, 100, 35		5.33 (1.45)	2.42 (1.33)
36, 35, 35		5.19 (1.43)	3.23 (1.47)
282, 35, 100		3.64 (1.59)	2.71 (1.48)
282, 100, 75		5.11 (1.45)	3.62 (1.44)
282, 100, 25		4.95 (1.55)	3.56 (1.48)
282, 35, 35		4.84 (1.5)	3.62 (1.34)
130, 35, 90		3.64 (1.27)	3.41 (1.41)
130, 100, 100		3.58 (1.54)	2.68 (1.39)
130, 75, 25		4.59 (1.58)	3.33 (1.35)
130, 35, 35		4.89 (1.56)	3.37 (1.36)
35, 40, 50		5.01 (1.29)	2.84 (1.44)
35, 80, 50		5.37 (1.42)	2.53 (1.23)
35, 80, 25		5.16 (1.63)	2.67 (1.48)
35, 40, 25		5.67 (1.28)	3.18 (1.47)
181, 35, 100		4.01 (1.55)	4.16 (1.47)
181, 100, 100		3.99 (1.62)	4.29 (1.56)
181, 100, 35		5.22 (1.36)	3.93 (1.18)
181, 35, 35		5.23 (1.33)	4.10 (1.33)
N.D., 0, 100		3.64 (1.58)	3.81 (1.61)
N.D., 0, 50		5.21 (1.39)	4.23 (1.14)
N.D., 0, 0		5.42 (1.60)	4.14 (1.61)

A.2.2.2. Regression Analysis

Table A.7.: *Summary of Multiple Linear Regression Analysis for Color Variables Predicting Image Schema Ratings (N=73)*

Image Schema	Variable	B	SE_B	β	R^2	adj. R_2	$F(9,21)$
BIG-SMALL	Intercept	-0.220	.157		.833***	.736	8.600
	Yellow	0.243	.148	.203			
	Blue	-0.017	.148	-.014			
	Cyan	0.267	.149	.223			
	Green	0.226	.149	.189			
	Orange	0.058	.148	.0649			
	Violet	0.207	.148	.172			
	Black	-0.810	.262	-.357**			
	Gray	0.129	.250	-.057			
	White	-0.762	.253	-.335**			
	Saturation	-0.006	.001	-.497***			
	Brightness	0.007	.001	.617***			
BRIGHT-DARK	Intercept	1.794	.384		.859***	.777	10.495
	Yellow	0.116	.362	.036			
	Blue	0.353	.362	.111			
	Cyan	0.022	.363	.007			
	Green	-0.252	.363	-.079			
	Orange	0.228	.362	.072			
	Violet	0.493	.362	.155			
	Black	0.160	.640	.0827			
	Gray	0.457	.610	.076			
	White	-0.353	.619	-.059			
	Saturation	0.001	.003	.033			
	Brightness	-0.028	.003	-.874***			
CLEAN-DIRTY	Intercept	-1.168	.278		.917***	.869	19.058
	Yellow	-0.402	.262	.212			
	Blue	0.070	.262	-.081			
	Cyan	0.175	.262	-.048			
	Green	0.096	.263	.021			
	Orange	-0.436	.262	.236*			
	Violet	0.007	.262	-.017			
	Black	1.063	.463	-.179			
	Gray	-0.245	.442	.106			
	White	0.534	.448	-.193			
	Saturation	0.003	.002	-.099			
	Brightness	-0.023	.002	-.847***			

Table A.7.: *Summary of Multiple Linear Regression Analysis for Color Variables Predicting Image Schema Ratings (N=73) (Continued)*

Image Schema	Variable	B	SE_B	β	R^2	adj. R^2	$F(9,21)$
FAR-NEAR	Intercept	-0.513	.129		.855***	.771	10.196
	Yellow	-0.157	.122	-.148			
	Blue	-0.552	.122	-.521***			
	Cyan	-0.359	.122	-.339**			
	Green	-0.115	.122	-.109			
	Orange	-0.241	.122	-.227			
	Violet	-0.267	.122	-.252*			
	Black	.148	.216	.074			
	Gray	-.438	.206	-.218*			
	White	-.200	.209	-.099			
	Saturation	0.004	.001	.429**			
	Brightness	0.008	.001	.733***			
FAST-SLOW	Intercept	1.154	.193		.931***	.891	23.312
	Yellow	0.086	.158	.046			
	Blue	-0.119	.149	-.064			
	Cyan	-0.194	.149	-.103			
	Green	-0.125	.149	-.066			
	Orange	0.331	.149	.176*			
	Violet	0.183	.149	.098			
	Black	-1.172	.263	-.330***			
	Gray	-0.107	.251	-.030			
	White	-0.617	.254	-.174*			
	Saturation	-0.004	.001	-.223**			
	Brightness	-0.018	.001	-.938***			
FULL-EMPTY	Intercept	-0.960	.207		.783***	.657	6.234
	Yellow	0.293	.223	.186			
	Blue	0.502	.206	.319*			
	Cyan	0.606	.207	.385**			
	Green	0.303	.205	.192			
	Orange	0.188	.223	.119			
	Violet	0.162	.207	.103			
	Black	-0.071	.376	-.024			
	Gray	0.765	.381	.256			
	White	-0.360	.381	-.121			
	Saturation	-0.007	.002	-.472**			
	Brightness	0.010	.002	.623***			

Table A.7.: *Summary of Multiple Linear Regression Analysis for Color Variables Predicting Image Schema Ratings (N=73) (Continued)*

Image Schema	Variable	B	SE_B	β	R^2	adj. R^2	F(9,21)
HARD-SOFT	Intercept	-0.794	.218		.840***	.747	9.063
	Yellow	0.264	.206	.156			
	Blue	-0.201	.206	-.060			
	Cyan	0.114	.206	.067			
	Green	0.364	.206	.214			
	Orange	0.195	.206	.115			
	Violet	0.011	.206	.007			
	Black	-0.760	.364	-.236*			
	Gray	-0.456	.347	-.141*			
	White	-0.071	.351	-.022			
	Saturation	-0.005	.002	-.329**			
	Brightness	0.013	.002	.729***			
HEAVY-LIGHT	Intercept	0.092	.493		.692**	.514	3.884
	Yellow	-0.487	.466	-.176			
	Blue	-1.086	.466	-.392*			
	Cyan	-0.314	.466	-.113			
	Green	-0.394	.466	-.142			
	Orange	-0.723	.466	-.261			
	Violet	-0.775	.465	-.280			
	Black	-1.136	.2465	-.216			
	Gray	-0.865	.822	-.165			
	White	-0.219	.795	-.042			
	Saturation	-0.007	.004	-.289			
	Brightness	0.019	.004	.662***			
LOUD-SILENT	Intercept	0.701	.244		.763**	.626	5.573
	Yellow	0.130	.269	.076			
	Blue	0.867	.254	.503**			
	Cyan	0.763	.254	.443**			
	Green	0.455	.254	.264			
	Orange	0.300	.254	.174			
	Violet	0.264	.254	.153			
	Black	-0.256	.448	-.078			
	Gray	0.403	.428	-.123			
	White	1.048	.434	.321*			
	Saturation	-0.007	.002	-.443**			
	Brightness	-0.011	.002	-.614***			

Table A.7.: *Summary of Multiple Linear Regression Analysis for Color Variables Predicting Image Schema Ratings (N=73) (Continued)*

Image Schema	Variable	B	SE_{B}	β	R^2	adj. R^2	$F(9,21)$
OLD-YOUNG	Intercept	-1.359	.195		.944***	.911	28.948
	Yellow	-0.371	.185	-.145			
	Blue	-0.169	.185	-.066			
	Cyan	-0.018	.185	-.007			
	Green	0.129	.185	.050			
	Orange	-0.501	.185	-.195*			
	Violet	-0.416	.185	-.162			
	Black	1.428	.326	.293***			
	Gray	-0.226	.311	-.046			
	White	-0.030	.315	-.006			
	Saturation	0.003	.002	.118			
	Brightness	0.026	.002	.998***			
PAINFUL-NOT PAINFUL	Intercept	0.513	.280		.577*	.332	2.357
	Yellow	0.436	.264	.037			
	Blue	0.524	.264	.329			
	Cyan	0.657	.264	.264*			
	Green	0.588	.264	.078*			
	Orange	0.493	.264	.138			
	Violet	0.263	.264	.259			
	Black	-0.564	.467	-.293			
	Gray	-0.105	.445	-.081			
	White	0.494	.451	.022			
	Saturation	-0.007	.002	-.439**			
	Brightness	0.002	.002	.223			
SMELLS GOOD-SMELLS BAD	Intercept	0.989	.175		.872***	.798	11.778
	Yellow	0.436	.166	.285*			
	Blue	0.292	.166	-.191			
	Cyan	0.282	.166	.184			
	Green	0.133	.166	.087			
	Orange	0.386	.166	.252*			
	Violet	0.197	.166	.129			
	Black	-0.531	.293	-.183			
	Gray	0.270	.279	.093			
	White	0.314	.283	.108			
	Saturation	-0.002	.001	-.133			
	Brightness	-0.015	.001	-.972***			

Table A.7.: *Summary of Multiple Linear Regression Analysis for Color Variables Predicting Image Schema Ratings (N=73) (Continued)*

Image Schema	Variable	B	SE_B	β	R^2	adj. R^2	F(9,21)
SMOOTH-ROUGH	Intercept	0.489	.171		.876***	.804	12.152
	Yellow	0.018	.161	.012			
	Blue	-0.188	.161	-.125			
	Cyan	-0.418	.161	-.277*			
	Green	-0.258	.161	-.171			
	Orange	0.033	.161	.022			
	Violet	-0.113	.161	-.075			
	Black	-0.453	.285	-.158			
	Gray	0.124	.271	.043			
	White	-0.539	.275	-.189			
	Saturation	0.002	.001	.123			
	Brightness	-0.013	.001	-.851***			
STRONG-WEAK	Intercept	-0.456	.268		.768***	.633	5.709
	Yellow	0.335	.254	.193			
	Blue	0.241	.254	.139			
	Cyan	0.631	.254	.364*			
	Green	0.330	.254	.190			
	Orange	0.219	.254	.126			
	Violet	0.147	.253	.085			
	Black	-1.172	.448	-.356*			
	Gray	0.122	.427	.037			
	White	-0.536	.433	-.163			
	Saturation	-0.012	.002	-.712***			
	Brightness	0.007	.002	.405**			
TASTES GOOD-TASTES BAD	Intercept	1.074	.208		.851***	.765	9.890
	Yellow	0.368	.196	.219			
	Blue	0.623	.196	.371**			
	Cyan	0.606	.196	.361**			
	Green	0.238	.196	.142			
	Orange	0.299	.196	.178			
	Violet	0.491	.196	.292*			
	Black	0.142	.346	.045			
	Gray	0.747	.330	.235*			
	White	0.440	.335	.138			
	Saturation	-0.000	.002	.001			
	Brightness	-0.015	.002	-.876***			

Table A.7.: *Summary of Multiple Linear Regression Analysis for Color Variables Predicting Image Schema Ratings (N=73) (Continued)*

Image Schema	Variable	B	SE_B	β	R^2	adj. R^2	F(9,21)
WARM-COLD	Intercept	-0.582	.249		.856***	.772	10.236
	Yellow	0.538	.235	.263*			
	Blue	1.776	.235	.869***			
	Cyan	1.522	.236	.744***			
	Green	0.856	.236	.419**			
	Orange	0.219	.235	.107			
	Violet	0.966	.235	.473**			
	Black	0.940	.416	.242*			
	Gray	1.331	.397	.343**			
	White	1.379	.402	.355**			
	Saturation	-0.002	.002	-.116			
	Brightness	-0.007	.002	-.318**			

Note. B = unstandardized regression coefficient; SE_B = Standard error of the coefficient; β = standardized coefficient

A.2.2.3. Regression Models

Table A.8.: *Color Regression Models of 16 Bipolar Image Schema Ratings (Japanese Subjects)*

Image Schema	Intercept (Red)	Hue	Saturation (S)	Brightness (B)
BIG-SMALL***	- 0.220	+ 0.243 (Yellow) - 0.017 (Blue) + 0.267(Cyan) + 0.226 (Green) + 0.058 (Orange) + 0.207 (Violet) - 0.810 (Black)** + 0.129 (Gray) -0.762 (White)**	- 0.006*S***	+ 0.007* B***
BRIGHT-DARK***	1.794	+ 0.116 (Yellow) + 0.353 (Blue) + 0.022 (Cyan) - 0.252 (Green) + 0.228 (Orange) +0.493 (Violet) + 0.160 (Black) + 0.457 (Gray) - 0.353 (White)	+ 0.001* S	- 0.028* B***
CLEAN-DIRTY***	1.168	+ 0.402 (Yellow) - 0.070 (Blue) - 0.175 (Cyan) + 0.096 (Green) + 0.436 (Orange)* - 0.007 (Violet) - 1.063 (Black) + 0.245 (Gray) - 0.534 (White)	- 0.003*S	- 0.023*B***
FAR-NEAR***	- 0.513	- 0.157 (Yellow) - 0.552 (Blue)*** - 0.359 (Cyan)** - 0.115 (Green) - 0.241 (Orange) - 0.267 (Violet)* + 0.148 (Black) - 0.438 (Gray)* - 0.200 (White)	+ 0.004*S**	+ 0.008*B***

Table A.8.: *Color Regression Models of 16 Bipolar Image Schema Ratings (Japanese Subjects) (Continued)*

Image Schema	Intercept (Red)	Hue	Saturation (S)	Brightness (B)
FAST-SLOW***	1.154	+ 0.086 (Yellow) +- 0.119 (Blue) - 0.194 (Cyan) - 0.125 (Green) + 0.331 (Orange) + 0.183 (Violet) - 1.172 (Black)*** - 0.107 (Gray) - 0.617 (White)*	- 0.004*S**	- 0.018*B***
FULL-EMPTY***	- 0.960	+ 0.293 (Yellow) + 0.502 (Blue)* + 0.606 (Cyan)** + 0.303 (Green) + 0.188 (Orange) + 0.162 (Violet) - 0.071 (Black) + 0.765 (Gray) - 0.360 (White)	- 0.007*S**	+ 0.010*B***
HARD-SOFT***	- 0.794	+ 0.264 (Yellow) - 0.201 (Blue) + 0.114 (Cyan) + 0.364 (Green) + 0.195 (Orange) + 0.011 (Violet) - 0.760 (Black)* - 0.456 (Gray)* - 0.071 (White)	- 0.005*S**	+ 0.013*B***
HEAVY-LIGHT**	0.701	- 0.487 (Yellow) - 1.086 (Blue)* - 0.314 (Cyan) - 0.394 (Green) - 0.723 (Orange) - 0.775 (Violet) - 1.136 (Black) - 0.865 (Gray) - 0.219 (White)	- 0.007*S	+ 0.019*B***

Table A.8.: *Color Regression Models of 16 Bipolar Image Schema Ratings (Japanese Subjects) (Continued)*

Image Schema	Intercept (Red)	Hue	Saturation (S)	Brightness (B)
LOUD-SILENT**	0.701	+ 0.130 (Yellow) + 0.867 (Blue)** + 0.763 (Cyan)** + 0.455 (Green) + 0.300 (Orange) + 0.264 (Violet) - 0.256 (Black) - 0.403 (Gray) + 1.048 (White)*	- 0.007*S**	- 0.011*B***
OLD-YOUNG***	- 1.359	- 0.371 (Yellow) - 0.169 (Blue) - 0.018 (Cyan) + 0.129 (Green) - 0.501 (Orange)* - 0.416 (Violet) + 1.428 (Black)*** - 0.226 (Gray) - 0.030 (White)	+ 0.003*S	+ 0.026*B***
PAINFUL-NOT PAINFUL*	0.513	+ 0.436 (Yellow) + 0.524 (Blue) + 0.657 (Cyan)* + 0.588 (Green)* + 0.493 (Orange) + 0.263 (Violet) - 0.564 (Black) - 0.105 (Gray) + 0.494 (White)	- 0.007*S**	+ 0.002*B
SMELLS GOOD-SMELLS BAD***	0.989	+ 0.436 (Yellow)* + 0.292 (Blue) + 0.282 (Cyan) + 0.133 (Green) + 0.386 (Orange)* + 0.197 (Violet) - 0.531 (Black) + 0.270 (Gray) + 0.314 (White)	- 0.002*S	- 0.015*B***

Table A.8.: *Color Regression Models of 16 Bipolar Image Schema Ratings (Japanese Subjects) (Continued)*

Image Schema	Intercept (Red)	Hue	Saturation (S)	Brightness (B)
SMOOTH-ROUGH***	0.489	+ 0.018 (Yellow) - 0.188 (Blue) - 0.418 (Cyan)* - 0.258 (Green) + 0.033 (Orange) - 0.113 (Violet) - 0.453 (Black) + 0.124 (Gray) - 0.539 (White)	+ 0.002*S	- 0.013*B***
STRONG-WEAK***	-0.456	+ 0.335 (Yellow) + 0.241 (Blue) + 0.631 (Cyan)* + 0.330 (Green) + 0.219 (Orange) + 0.147 (Violet) - 1.172 (Black)* + 0.122 (Gray) - 0.536 (White)	- 0.012*S***	+ 0.007*B**
TASTES GOOD-TASTES BAD***	1.074	+ 0.368 (Yellow) + 0.623 (Blue)** + 0.606 (Cyan)** + 0.238 (Green) + 0.299 (Orange) + 0.491 (Violet)* + 0.142 (Black) + 0.747 (Gray)* + 0.440 (White)	- 0.000*S	- 0.015*B***
WARM-COLD***	- 0.582	+ 0.528 (Yellow)* + 1.776 (Blue)*** + 1.522 (Cyan)*** + 0.856 (Green)** + 0.219 (Orange) + 0.966 (Violet)** + 0.940 (Black)* + 1.331 (Gray)** + 1.379 (White)**	- 0.002*S	- 0.007*B**

A.3. Study 3: Confirmation of Color-Image Schema Associations with German Subjects

Instructions (in German)

Figure A.1.: Instruction screen.

A.4. Study 4: Confirmation of Color-Image Schema Associations with Japanese Subjects

Instructions (in Japanese)

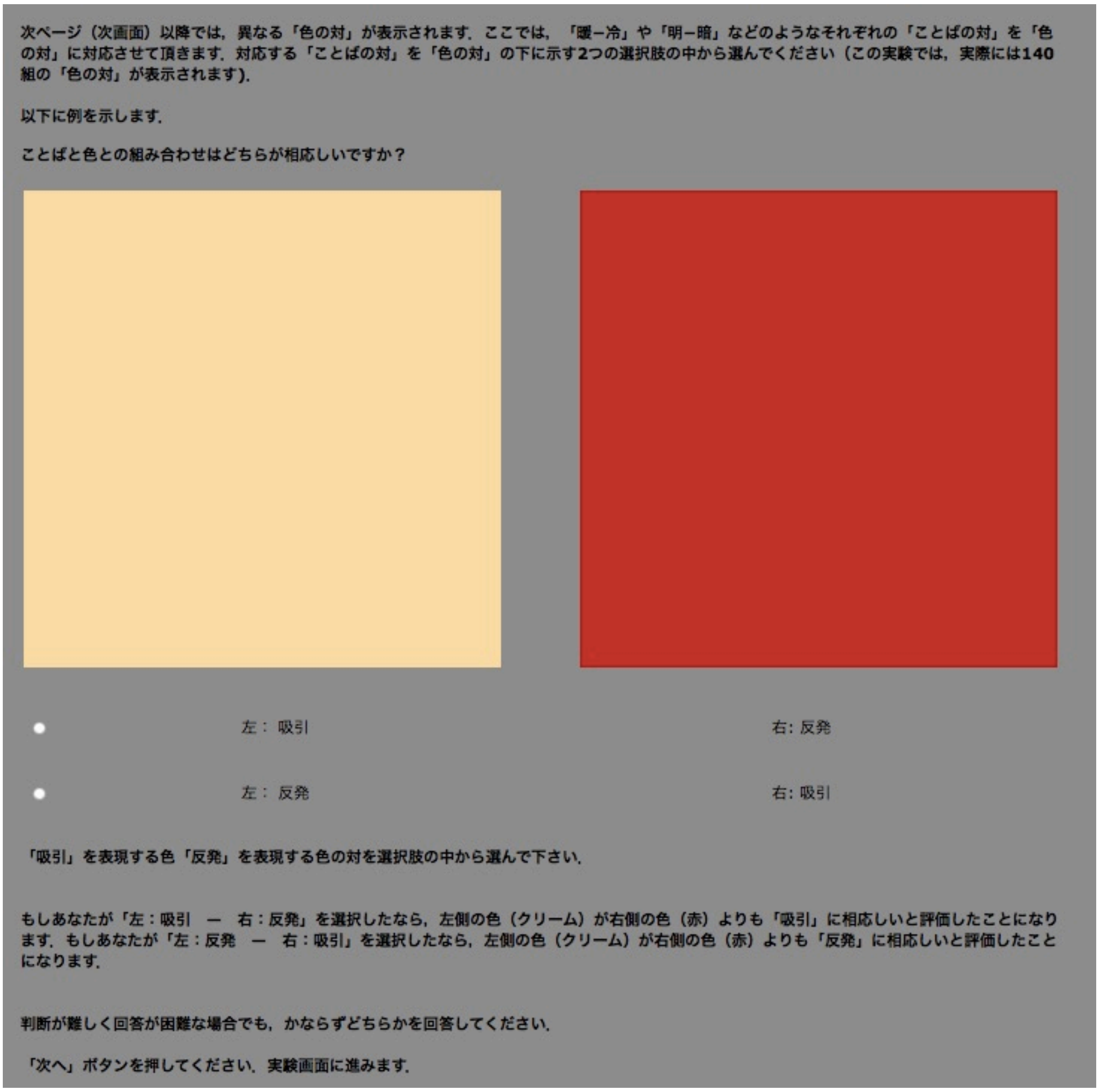

Figure A.2.: Instruction screen.

A.5. Study 5: Conveying Abstract Information with Color

A.5.1. Preference questionnaire (translated into English)

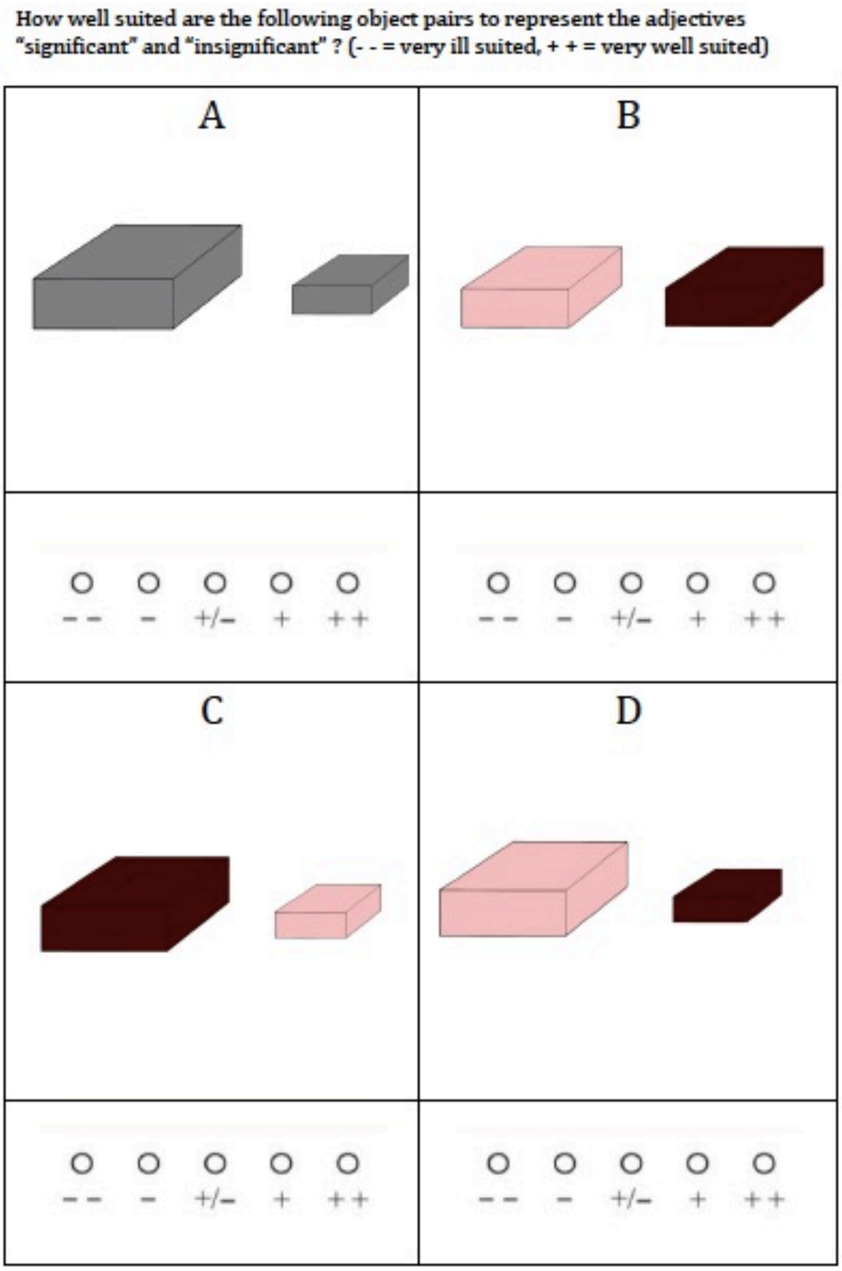

Figure A.3.: Sample questionnaire version for BIG-SMALL.

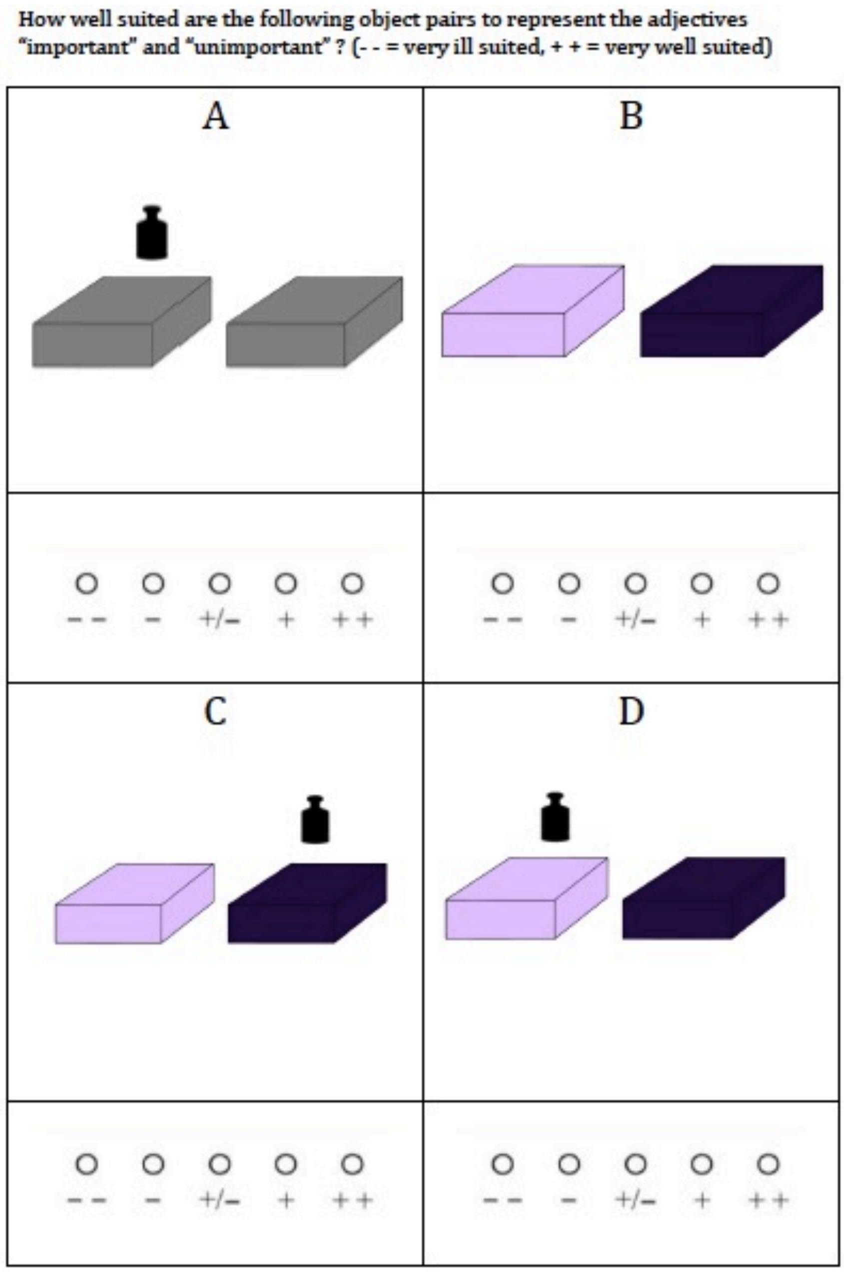

Figure A.4.: Sample questionnaire version for HEAVY-LIGHT.

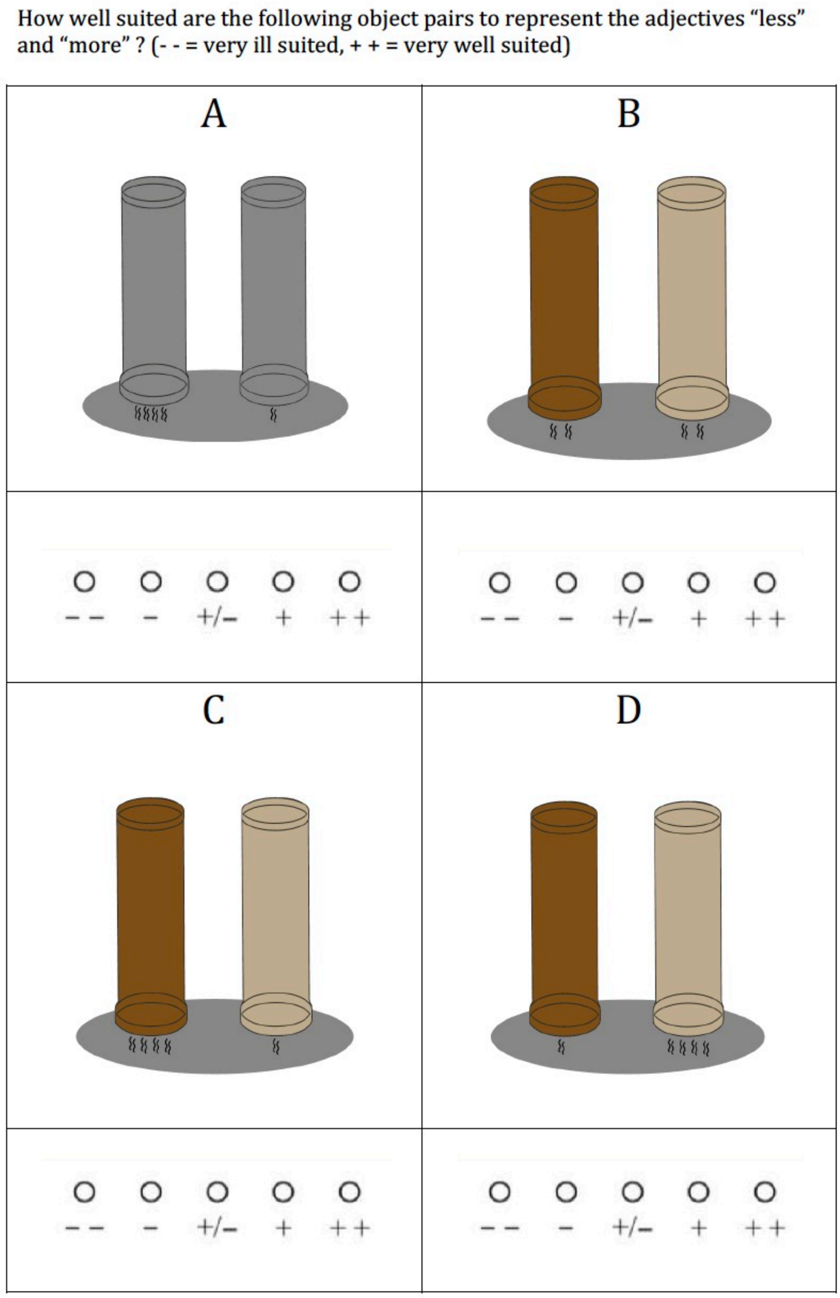

Figure A.5.: Sample questionnaire version for STRONG-WEAK.

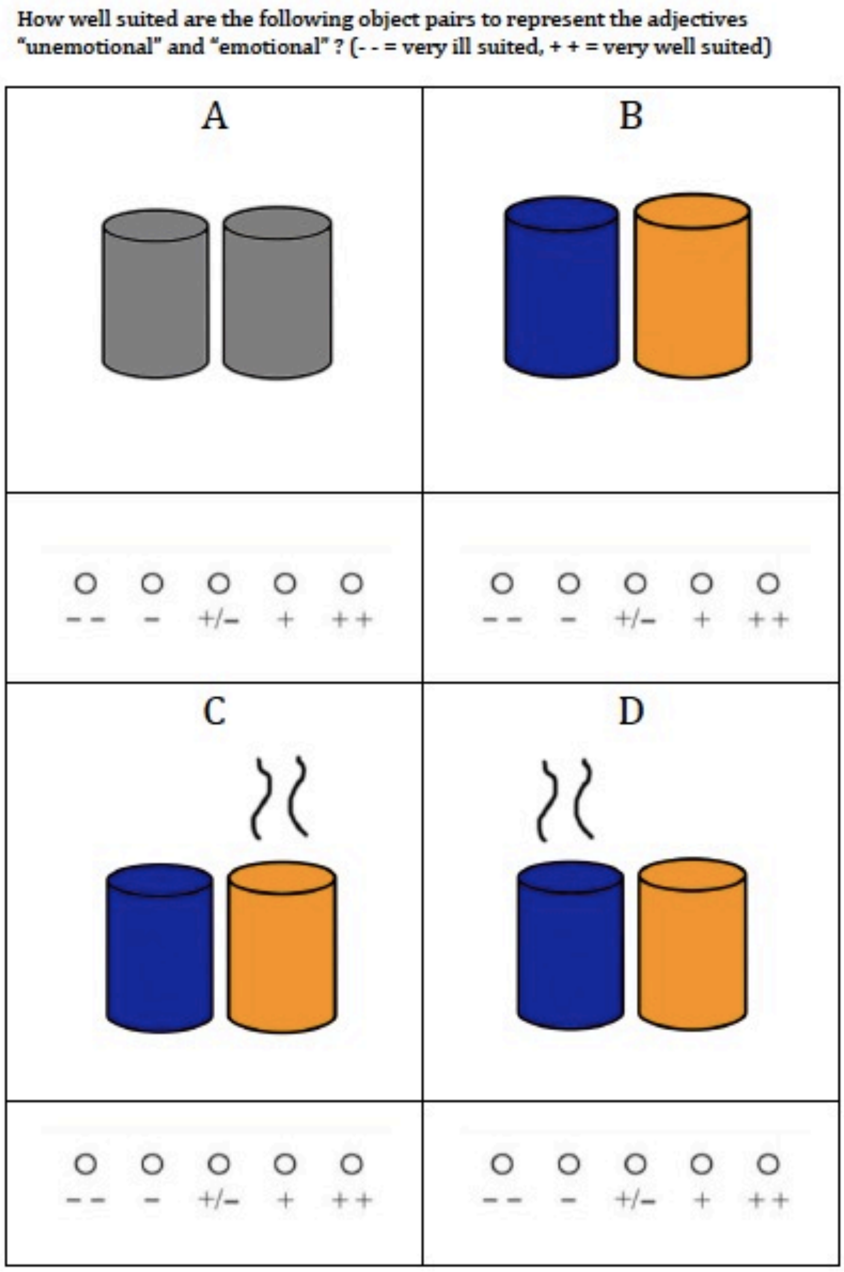

Figure A.6.: Sample questionnaire version for WARM-COLD.

Table A.9.: *Metaphor Consistency in the Answers Across Four Presentation Styles in Six Conditions (28 German Subjects)*

	Haptic		Color		Congruence		Incongruence		Blindfolded		Looking	
	%	ϰ	%	ϰ	%	ϰ	%	ϰ	%	ϰ	%	ϰ
Metaphor (... is big - ... is small)	**93**	**0.86**	**83**	**0.78**	**89**	**0.78**	**88**	**0.76**	**93**	**0.86**	**78**	**0.56**
significant - insignificant	93	0.86	86	0.72	85	0.70	89	0.78	100	1.00	68	0.36
important - unimportant	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
powerful - powerless	93	0.86	82	0.64	89	0.78	93	0.86	93	0.86	89	0.78
more - less	100	1.00	89	0.78	96	0.92	96	0.92	96	0.92	89	0.78
knowing - unknowing	86	0.72	75	0.50	82	0.64	71	0.42	82	0.64	64	0.28
Metaphor (... is heavy - ... is light)	**94**	**0.88**	**76**	**0.52**	**95**	**0.90**	**74**	**0.48**	**95**	**0.90**	**76**	**0.52**
important - unimportant	89	0.78	64	0.28	93	0.86	86	0.72	100	1.00	64	0.28
sad - happy	89	0.78	100	1.00	89	0.78	25	-0.50	93	0.86	100	1.00
more - less	100	1.00	69	0.38	89	0.78	93	0.86	96	0.92	62	0.24
difficult - easy	96	0.92	71	0.42	100	1.00	86	0.72	93	0.86	79	0.58
guilty - innocent	96	0.92	92	0.84	100	1.00	79	0.58	93	0.86	92	0.84
Metaphor (... is strong - ... is weak)	**85**	**0.70**	**84**	**0.68**	**90**	**0.80**	**72**	**0.44**	**81**	**0.62**	**81**	**0.62**
more - less	86	0.72	93	0.86	96	0.92	82	0.64	89	0.78	89	0.78
powerful - powerless	86	0.72	82	0.64	89	0.78	82	0.64	79	0.58	89	0.78
intense - less intense	89	0.78	75	0.50	100	1.00	92	0.84	86	0.72	71	0.42
competent - incompetent	86	0.72	79	0.58	79	0.58	75	0.50	68	0.36	64	0.28
male - female	79	0.58	93	0.86	93	0.86	43	-0.14	82	0.64	93	0.86
Metaphor (... is warm - ... is cold)	**88**	**0.76**	**88**	**0.76**	**90**	**0.80**	**80**	**0.60**	**84**	**0.68**	**80**	**0.60**
intimate - distant	100	1.00	89	0.78	93	0.86	80	0.60	100	1.00	89	0.78
problematic - unproblematic	64	0.28	53	0.06	71	0.42	69	0.38	68	0.36	43	-0.14
emotional - unemotional	96	0.92	96	0.92	100	1.00	n/a	n/a	96	0.92	82	0.64
active - inactive	96	0.92	89	0.78	93	0.86	87	0.74	85	0.70	86	0.72
happy - sad	82	0.64	93	0.86	93	0.86	85	0.70	75	0.50	100	1.00
Overall	**90**	**0.80**	**83**	**0.78**	**91**	**0.82**	**79**	**0.58**	**88**	**0.76**	**79**	**0.58**

Table A.10.: *Metaphor Consistency in the Answers Across Four Presentation Styles in Six Conditions (47 Japanese Subjects)*

	Haptic		Color		Congruence		Incongruence		Blindfolded		Looking	
	%	$\varkappa$	%	$\varkappa$	%	$\varkappa$	%	$\varkappa$	%	$\varkappa$	%	$\varkappa$
Metaphor (... is big - ... is small)	**81**	**0.62**	**88**	**0.76**	**89**	**0.78**	**51**	**0.02**	**80**	**0.60**	**81**	**0.62**
significant - insignificant	82	0.64	98	0.96	91	0.82	43	-0.14	80	0.60	93	0.86
important - unimportant	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
powerful - powerless	74	0.48	95	0.90	98	0.96	36	-0.28	86	0.72	93	0.86
more - less	98	0.96	84	0.68	100	1.00	80	0.60	82	0.64	68	0.38
knowing - unknowing	66	0.32	75	0.50	66	0.32	43	-0.14	64	0.28	59	0.18
Metaphor (... is heavy - ... is light)	**95**	**0.90**	**90**	**0.80**	**97**	**0.94**	**68**	**0.36**	**89**	**0.78**	**90**	**0.90**
important - unimportant	100	1.00	86	0.72	100	1.00	66	0.32	95	0.90	93	0.86
sad - happy	77	0.54	89	0.78	89	0.78	45	-0.10	64	0.28	91	0.82
more - less	100	1.00	84	0.68	98	0.96	98	0.96	95	0.90	86	0.72
difficult - easy	98	0.96	98	0.96	98	0.96	70	0.40	100	1.00	93	0.86
guilty - innocent	98	0.96	91	0.82	100	1.00	59	0.18	95	0.90	84	0.68
Metaphor (... is strong - ... is weak)	**95**	**0.90**	**87**	**0.74**	**93**	**0.86**	**75**	**0.50**	**87**	**0.74**	**91**	**0.82**
more - less	100	1.00	89	0.78	95	0.90	73	0.46	93	0.86	98	0.96
powerful - powerless	95	0.90	79	0.58	95	0.90	88	0.76	93	0.86	98	0.96
intense - less intense	86	0.72	84	0.68	89	0.78	76	0.52	73	0.46	91	0.82
competent - incompetent	93	0.86	77	0.54	89	0.78	77	0.54	84	0.68	73	0.46
male - female	98	0.96	95	0.90	98	0.96	68	0.36	91	0.82	98	0.96
Metaphor (... is warm - ... is cold)	**88**	**0.76**	**87**	**0.74**	**89**	**0.78**	**58**	**0.16**	**86**	**0.72**	**85**	**0.70**
intimate - distant	91	0.82	95	0.90	93	0.86	57	0.14	89	0.78	93	0.86
problematic - unproblematic	55	0.10	45	-0.10	55	0.10	59	0.18	61	0.22	32	-0.36
emotional - unemotional	98	0.96	98	0.96	98	0.96	61	0.22	100	1.00	100	1.00
active - inactive	98	0.96	100	1.00	100	1.00	52	0.04	86	0.72	98	0.96
happy - sad	100	1.00	98	0.96	100	1.00	61	0.22	95	0.90	100	1.00
Overall	**89**	**0.78**	**88**	**0.76**	**92**	**0.84**	**63**	**0.26**	**86**	**0.72**	**86**	**0.72**

A.5.2. Results - Effectiveness

A.6. Study 6 - Semantic Stroop Effect in Japanese Subjects

Instructions (in Japanese)

実験への御協力ありがとうございます．まず中央の注視点を見ていてください．その後，ある単語が提示されますので，その単語の色を答えて頂きます．できるだけ早く，適切だと思う色のキーを押してください．できるだけ間違えないようにお願いします！なお，利き手の人差し指のみで操作して下さい．

はじめに，実験のための練習をします．なお，練習セッションでは，正解か不正解かをフィードバックします．何かキーを押してください.

A.7. Study 7 - Semantic Stroop Effect in German Subjects

Instructions (in German)

Willkommen zum zweiten Teil der Studie.

Im nachfolgenden sehen Sie immer zunächst einen Fixationspunkt in der Mitte des Bildschirmes. Danach folgt ein farbig geschriebenes Wort. Ihre Aufgabe ist es, die Art des Wortes zu klassifizieren. Dazu drücken Sie die entsprechende Taste auf der Tastatur (H oder J). Die Zuordnung der Kategorie zur Antworttaste wird Ihnen immer am unteren Bildschirmrand angezeigt.

Versuchen Sie, so schnell wie möglich zu reagieren. Wichtiger ist jedoch, dass Sie versuchen, so wenig Fehler wie möglich zu machen. Reagieren Sie nur mit dem Zeigefinger Ihrer dominanten Hand. Bitte vermeiden Sie, auf die Windows-Taste zu drücken, da dies zum Absturz des Programms führen kann.

Bitte drücken Sie eine beliebige Taste.

A.8. Study 8 - Reverse Semantic Stroop Effect in German Subjects

A.8.1. Instructions Practice Block (in German)

Es folgt ein Übungsblock.

In diesem Übungsblock erhalten Sie Rückmeldung, ob Sie richtig oder falsch reagiert haben. Bitte ordnen Sie jedes Wort einer der folgenden Kategorien zu:

Antworttaste H = negativ

Antworttaste J = positiv

Bitte reagieren Sie nur mit dem Zeigefinger Ihrer dominanten Hand.

Sie starten den Übungsblock mit einer beliebigen Taste.

A.8.2. Instructions Main Experiment (in German)

Jetzt beginnt das eigentliche Experiment. Sie bearbeiten im folgenden 8 Blöcke ähnlich des Übungsdurchgangs.

Ihre Aufgabe ist es wieder, die Wörter zu klassifizieren. Die Kategorien ändern sich von Block zu Block. Jeder Block dauert ca. 3 Minuten. Nach jedem Block können Sie eine Pause machen. Bitte machen Sie keine Pause während des Blocks. Sie erhalten ab sofort kein Feedback mehr, versuchen Sie so wenig Fehler wie möglich zu machen. Reagieren Sie nur mit dem Zeigefinger Ihrer dominanten Hand.

Sie starten den ersten Experimentalblock mit einer beliebigen Taste.

Bitte ordnen Sie jedes Wort einer der folgenden Kategorien zu:

Antworttaste H = kalt (intim/passiv/emotional)

Antworttaste J = warm (distanziert/aktiv/sachlich)

Bitte reagieren Sie nur mit dem Zeigefinger Ihrer dominanten Hand.

Sie starten den Block mit einer beliebigen Taste.