DTU Library

Improving Visual Information Search in Emergencies through Using Cognitive Theory in Design

Andersen, Emil

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

Improving Visual Information Search in Emergencies through Using Cognitive Theory in Design

Emil Andersen

PhD Thesis

November, 2018

Engineering Systems Division
DTU Management Engineering
Technical University of Denmark

Title:	**Improving Visual Information Search in Emergencies through Using Cognitive Theory in Design**
Type:	PhD thesis
Date:	30-11-2018
Author:	Emil Andersen
Supervisors:	Anja Maier Igor Kozin Stine Strand
University:	Technical University of Denmark
Department:	DTU Management Engineering
Division:	Engineering Systems
Address:	Produktionstorvet Building 358 DK-2800 Kgs. Lyngby www.man.dtu.dk
Cover picture:	Colourbox.com

Acknowledgements

My first and greatest thank you goes to my three supervisors Professor Anja Maier, Dr. Igor Kozine and Stine Strand for their help and support during my project. Thank you for broadening my horizon and for all you have taught me through my years at DTU and during my visits in Halden. And thank you for giving me the opportunity and guidance to travel from the lab into the 'real world'.

I further would like to recognise the funding agents for this PhD research: The OECD Halden Reactor Project and the Department of Management Engineering of the Technical University of Denmark. In particular, thank you to the Engineering Systems Division of DTU Management Engineering and the Institute for Energy Technology's Man, Technology, Organisation Labs for hosting me during these last three years. Adding to this, I am grateful for our collaborators in Halden for making me feel welcome during my visits. In particular, I would like to thank Andreas Bye, Espen Nystad, Magnhild Kaarstad and Salvatore Massaiu.

A special thanks goes out to my colleagues at the Engineering Systems Division, in particular Lisbeth Bjerre Damgaard, who I could always count on for help when bureaucracy had me down, and my fellow PhD students, Verena Stingl, Sebastiano Piccolo, Julia Thorpe and Dr. Miroslava Tegeltija, who were with me through the duration of my project, and whose support, insight and friendship made every workday much more fun.

I would further like to thank my hosts and collaborators at the Integrated Design Innovation Group of Carnegie Mellon University – thank you for making my stay both exciting and productive. In particular, I would like to give my thanks to Dr. Kosa Goucher-Lambert and Professor Jonathan Cagan for facilitating my visit, for their great ideas and challenging questions during our collaboration, and for making me feel at home during my stay.

Finally, I would like to thank Professor John Paulin Hansen and Dr. Philip Cash for our coffee sparring sessions.

Abstract

In emergencies, fast and accurate decision making is essential for avoiding adverse effects. Optimal performance hinges on the ability of operators to retrieve the information they need from their surroundings and to translate this information into action. However, in the stressful and complex environment of an emergency, performance can be compromised if search for information is slowed or hindered by few or inadequate visual aids and guides.

The present thesis represents an effort towards using cognitive theory to improve search for information through better visual design. The research conducted has as its basis that findings, and in some cases methodology, from highly controlled experiments from cognitive science are translated through literature review and original research to guidelines and concrete suggestions to be used by visual display and procedure designers in the real world.

In total, six papers are presented. The first paper reviews the literature in the field of design research pertaining to human behaviour, establishing the benefits of cognitive theory and methodology in design research. Papers two through five investigate, through literature review and original experiments, how findings from the field of experimental psychology on attention could be applied. In particular, the effect on attention of specific colours, cognitive and visual load, visual dilution, and distracting stimuli were investigated. The main findings were that specific colours have different effects on attention allocation, especially under higher visual load, and that the Feature Integration Theory of Treisman and Gelade was particularly successful in generalising its predictions to a more applied context. Paper six investigates, through literature review and re-analysis of behavioural patterns of active, experienced nuclear control room operators in a realistic nuclear power plant emergency scenario, whether literature from the decision making sciences can predict the behavioural biases of operators and whether such biases could be alleviated through design. It was found that variance in operator performance could in part have been explained caused by confirmation bias and the bias of misapplied expertise, which could potentially be used as the basis for interventions in procedure design.

Taken together, the research contributes novel findings to both design research and experimental psychology, and offers suggestions for how the findings can be applied to improve visual search in emergencies.

Publications

This thesis is based on the following papers:

- Andersen, E., Piccolo, S., Maier, A. (Working Paper): Improving Deductive Studies of Human Behaviour in Design Through Cognitive Science – A Review and Look Forward. To be submitted to Design Science.
- Andersen, E., & Maier, A. (2016). What Captures Gaze in Visual Design? Insights from Cognitive Psychology. DS 85-1: Proceedings of NordDesign 2016, Volume 1, Trondheim, Norway, 10th-12th August 2016.
- Andersen, E., & Maier, A. (2017). The attentional capture of colour in visual interface design: a controlled-environment study. In DS 87-8 Proceedings of the 21st International Conference on Engineering Design (ICED 17) Vol 8: Human Behaviour in Design, Vancouver, Canada, 21-25.08. 2017.
- Andersen, E., & Maier A. (Submitted): The Attentional Guidance of Colour in Increasingly Complex Displays. Submitted to Applied Ergonomics.
- Andersen, E., Goucher-Lambert, K., Cagan, J., Maier, A. (Submitted): Attention Affordances: Applying Attention Theory in the Design of Complex Visual Interfaces. Submitted to the Journal of Experimental Psychology: Applied.
- Andersen, E., Kozin, I., & Maier, A. (2018). Biased Decision Making in Realistic Extra-Procedural Nuclear Control Room Scenarios. In *8th International Conference on Design Computing and Cognition*.

In addition, two Work Reports were submitted to the Halden Research Project:

- Andersen, E., Kozine, I., Maier, A., Massaiu, S.: Decision-making in control room emergencies: A re-analysis of a HAMMLAB experiment in the 'heuristics and biases' paradigm (2018). Submitted Work Report to the OECD Halden Reactor Project, see Appendix A.
- Andersen, E. & Maier, A.: Biased Information Search with Missing or Unreliable Information (2018).Submitted Work Report to the OECD Halden Reactor Project, see Appendix B.

Contents

CHAPTER 1:
Motivation, Methodological Approach and Research Question

1.1. Motivation

In high stakes or emergency situations, adequate and efficient decision making is essential for minimising detrimental effects or potentially devastating consequences. Optimal performance hinges on the ability of operators to retrieve the information they need from their surroundings and to translate the information to action. However, in the stressful and complex environment of an emergency, performance can be compromised if there are few or inadequate visual aids and guides available (Fernandes & Braarud, 2015).

A basic assumption of much of common sense thinking about human decision making is that information provided will somewhat directly predict behaviour. For example, one might assume that a decision maker uses all the information provided, especially if one has a guarantee that the person has read it. One may also assume that a decision maker considers in relative equal proportion all the information presented with equal or adequate degrees or weights of importance. However, research has now shown that this is seldom the case. Instead decision makers seem to follow consistent, irrational rules of both the way information is found and interpreted and the way that decisions are made in relation to these interpretations (Tversky & Kahneman, 1974). Furthermore, their attention will be captured by irrelevant stimuli, leading to slow and inefficient search (Theeuwes, 2010; Wolfe & Horowitz, 2004). Within the cognitive sciences, researchers have developed theories that predict the attentional allocation and information-searching patterns of many different types of humans under many different conditions. While applying basic principles of human cognition has shown promise for improving visual design (Faerber & Carbon, 2013; Nørager, 2009), translating the insights from attention theory and decision science to actionable design guidelines that could aid in creating the visual interfaces used emergencies has seldom been attempted. Given that theories in these fields are almost exclusively deduced from highly controlled laboratory experiments, with only one or two functions being examined at a time, the findings and predictions from these theories are hard to translate directly to the complex situations of reality. As an unfortunate consequence, designers may rely on outdated or obsolete theories of the attention system in their work. Conversely, while it is assumed that the theories derived from these experiments are generalisable to everyday situations, in reality it has seldom been tested. Understanding the predictive powers of theories of attention in large-scale, complex situations could therefore provide valuable insights for academics in these fields as well.

The present PhD project attempts to alleviate this gap by translating and testing the theories of cognitive sciences to a visual design cases. The thesis takes as its starting place the practices of the visual designer by taking stock of what has currently been done, and then adds to it suggestions of which theories and methods from cognitive science could be beneficial for visual design. It is my hope that this thesis thereby contributes to both our understanding of the basic laws of human processing and the understanding of design practitioners on what should be done to better design for desired decisions.

1.2. Methodological Approach-

To achieve these research goals, this PhD is based on studies that employ a variety of complementary approaches that reflect the interplay between highly controlled experiments and realistic use cases: On the one hand, this PhD consists of papers that test specific interactions between the attention system and visual characteristics that have high relevance for designers, such as colour and salience, using experimental studies with student participants. On the other hand, this PhD includes a study of the behaviour of highly experienced, active nuclear control room operators in realistic emergency simulations. Each approach is elaborated below.

The overarching scope of this PhD project was to test theories from literature in an applied context. While some specific predictions were derived from the reviewed literature, these were seldom linked to novel theory building, but rather were constructed to reflect back on pre-existing theories. The project thus followed the structure of a Type 3 research project as outlined by the Design Research Methodology (Blessing & Chakrabarti, 2009, p. 18), wherein literature reviews are conducted and form the basis for a hypothesis/hypotheses, which is/are tested in a prescriptive study. In the broader scientific lens, this PhD project thus mostly falls in the theory testing part of the scientific method as popularised by Popper (Popper, 1959), which is illustrated in Figure 1: First, an observation is made about the world, second, scientific theory is made to explain the observation in a testable manner, third, predictions from the theories are tested, fourth, revisions are made to the theories based on the outcomes of the tests.

Figure 1: An illustration of the scientific process as popularised by Popper (Popper, 1959).

Elaborating on this based on a more recent outline of the scientific process (Cash, 2018, p. 88), this PhD project defines variables and limitations in both the fields of design and experimental psychology, proposes how the two fields can be combined, translates these relationships into predictions, and tests the predictions. Specifically, this was done through six studies described in the chapters and papers of this PhD thesis:

The first study, presented in chapter three, was a literature review of all deductive studies on human behaviour in seven prominent design research journals. The review was conducted by reading the abstracts of all papers published in the included journals since their inception (1979 for the oldest included paper) and finding all deductive studies on human behaviour for further review. The inclusion criteria are elaborated in chapter 2. Following this, the papers were evaluated with regards to the coherence of topics and methods and suggestions were made regarding how including cognitive theory could further build the field.

The second study, presented in chapter three, was similarly a literature review, but focused on outlining the factors that determine where gaze is directed, and how it can be used in design. The review was conducted by outlining the most prominent theories and well-described findings regarding the allocation of attention and making suggestions on their implication for both design research and design practice.

The third and fourth studies, presented in chapter three, describe deductive computerised experiments created with basis in attention theory and the paradigms used in this research tradition, but with added complexity in the displays to increase its ecological validity. Specifically, the paradigm used was designed to test the attentional guidance of six specific colours (red, blue, yellow, green, orange, and purple) through a target-detection task, wherein participants had to determine whether a target coloured circle was present amongst 19 other coloured circles. The target was present in two thirds of the trials, and each colour appeared as a target with each other possible combination of colours. Additionally, complexity varied across trials as there was either one, two, three or four possible distractor colours for any given trial. The third study (E. Andersen & Maier, 2017) investigated only the average response times to the specific colours, whereas the fourth study (Andersen & Maier, submitted) included a larger sample size (17 as opposed to 11), which allowed for investigating the interaction effect between attentional guidance of individual colours and the complexity of the display. The outcome variable was reaction time for accurate trials where the target was present.

The fifth study, presented in chapter three, describes another deductive computerised experiment created with basis in attention theory and the paradigms used in this research tradition. Again, the paradigm was designed to test theory found in attention theory in a more ecologically valid setting. In this experiment, the complexity (and therefore possibly ecological validity) of the design was increased further, in part through manipulating the cognitive load (high or low) through a memory task, in part through using a complex visual display designed to contain objects that resembled the ones found in displays of nuclear control room operators. In the study, student participants searched for three digit numbers (e.g. 194) that were presented in 'barometers' based on their knowledge of either a target number, or a target number and the colour of the barometer that contained the number. In addition to testing the effect of cognitive load and knowledge about the target colour, the experiment included tests derived from attention theory for five additional possible effects: 1) The attentional guidance or capture of specific colours (white, red, blue, yellow, green, purple, and orange), 2) the effect of visual load, measured by presenting the target number amongst similar numbers (524) or amongst three zeros (000), 3) the effect of irrelevant distractors in the form of 'trend graph' objects, 4) the effect of a central distractor that either shared the colour of the target barometer or the distractor barometers, 5) the effect of visual dilution, measured as the interaction effect of the distractor effect and the irrelevant distractor effect. The outcome variable was reaction time on accurate trials.

The sixth study, presented in chapter four, was a re-analysis of a dataset collected on a full-scale simulation study of nuclear power plant emergency, wherein experienced, active nuclear control room operators worked in realistic emergency scenarios at the Halden Man-Machine Laboratory (HAMMLAB). The re-analysis was done through an exploratory reading of performance ratings and comments on behaviour of a process expert and human factors expert as part of their original analysis of the data, as well as a re-visit of video and audio recordings. The study compared the behavioural patterns of the crews in relation to performance, and explored whether differences in these variables could in part be explained by biased behaviour due to confirmation bias or bias of misapplied expertise.

In sum, the present PhD thesis employed a variety of research methods, ranging from highly controlled experiments with student participants on the attentional guidance of colour to a re-analysis of data from a realistic simulation study with active, experienced nuclear control room operators. In all cases, however, the studies were grounded in using cognitive theory for cases that could have impact for improving design practice.

1.3. Research Question and Objectives

This PhD thesis had the following research question as its basis:

- How can visual design for emergencies be improved through improving design research [1], and through cognitive theory on attention [2], biases and heuristics [3]?

Answering the research question was approached through three angles that represent the three bracketed highlights in the research question. Each angle is elaborated and justified below:

1. Improving design research

Despite its relatively young age, the field of design research has provided a rich and diverse set of insights about designer practice, designers themselves, and the perception of designed objects. However, whether due to its young age or other reasons, the field has so far yet to establish comprehensive theories and best-practice methodologies as is seen in older fields (Papalambros et al., 2015; Cash, 2018). Conversely, the majority of cognitive theory is inseparable from the methods underlying it. Indeed, the field of psychology, which underlies much of cognitive theory, was founded with the objective of developing methods for systematically investigating the mind (Hergenhahn & Henley, 2014). Therefore, applying cognitive theory and methodology to design research provides an opportunity to potentially solve some of the fields' standing problems. As such, the first research angle of this thesis was to apply methodology and theory from cognitive science as a means of possibly improving design research on human behaviour, which in turn should lead to improved insights on visual design for emergencies.

2. *Applying attention theory to design*

The attention system is the primary determinant of which pieces of information enter consciousness (Baars & Franklin, 2007; Dehaene et al., 2014). It follows, that if one wishes to improve information search through visual design, then one should apply knowledge about the visual attention system. Experimental psychologists have developed both comprehensive theories of attention (e.g. Bundesen, 1990; Bundesen & Habekost, 2008; Wolfe et al., 1989) and predictions for specific conditions (e.g. under Load: Lavie et al., 2004; Lavie & Tsal, 1994) and types of stimuli (e.g. emotional stimuli: Öhman et al., 2001). However, these insights have thus far not been investigated within design research, and are seldom used in design practice. Conversely, the insights from attention theory are seldom tested in more complex, ecologically valid contexts. To alleviate these concerns, the second research angle of this PhD was to apply and test attention theory in complex situations that better resemble the use cases that designers encounter.

3. *Applying bias and heuristics literature to design*

While attention theory relates to the smaller scale allocation of attention when a visual scene is viewed, the larger scale information searching strategies and how information is evaluated once it enters consciousness are generally classified as relating to the decision making sciences. Research on decision making has revealed that both human information search and evaluation seldom follow a rational, consistent trajectory. Instead, both have been shown to follow predictably irrational patterns characterised by biased behaviour and decision making short-cuts – usually referred to as heuristics (Tversky & Kahneman, 1974). As with attention theory, applying insights about these heuristics and biases is essential for making design that fits the intended human users. However, as with attention theory, this has thus far seldom been explored in design research. Therefore, the third research angle of this PhD was to apply and test insights from the heuristics and bias literature to design relevant cases.

1.4. Thesis Structure

Having outlined the research question, as well as the three angles applied for answering it, the remainder of this thesis is structured in four parts: one for each angle and a concluding discussion based on the presented research. This thesis is paper based, consisting of five chapters. In chapter one the motivation, research question and methodological approach were introduced. In chapter two, the need for cognitive theory in design research is further established through a literature review of research on human behaviour in design. In chapter three, literature review and original research is presented that show how attention theory could be applied in visual design. In chapter four, literature review and original research is presented that show how decision making theory could be applied to understand expert decision making and improve design of procedures. In chapter five, conclusions and implications are discussed.

CHAPTER 2:

Cognitive Theory in Design

2.1. Chapter Introduction

While accumulating knowledge about human cognition could provide a richer understanding of how to create better visual designs for emergencies, the degree to which the implementation of such knowledge is successful inevitably depends on the designer making the actual visual artefacts. Research on Human Behaviour in Design encompasses both studies of designer behaviour and the behaviour of participants interacting with designed artefacts. The research in this field has thus created insights on both how to improve the design process and concrete suggestions that designers can apply when creating visual artefacts. Recently, however, concerns have been voiced within the field of design research in general, with several researchers calling for increased scientific rigour through stronger relation to theory and more rigorous and consistent use of methods (Cash, 2018; Dinar et al., 2015; Papalambros, 2015; Vasconcelos & Crilly, 2016). As the first angle towards improving visual design for emergencies, this chapter focuses on the field of design research and how it might be improved. Specifically, given that Human Behaviour in Design is closely related to the cognitive sciences due to their object of study, and given that this neighbouring field contain several comprehensive theories and rigorous methodologies, this chapter focuses on whether insights from the cognitive sciences could improve human behaviour in design research through enabling theory building and providing methodological insights.

2.2. Research and Findings

The paper presented in this chapter (Andersen et al., working paper) addresses this question through an extensive literature review of the deductive studies conducted within the topic of human behaviour in design. The review was conducted by examining all deductive studies on human behaviour published in seven prominent design journals, from their inception until 2017, and evaluating the topic and method coherence found across papers. Results showed that the field was characterised by a variety of studies that explored many different topics, rather than studying the various aspects of fewer topics. Furthermore, the results indicated that this diversity leads to the findings giving a broad overview of the possible areas of design, rather than relating to common theories or to a more comprehensive understanding of the topics. Based on these findings, the review provides suggestions for cognitive theories and methodological insights that could both further the understanding of the individual topics, and serve as a basis for common theories across the diverse topics. While the review offers suggestions for several different directions of human behaviour in design research due to the nature of the review material, the present thesis has as its foundation the observation that few studies on perception of designed objects employed methods or theory from cognitive science to guide their studies. The aim of this thesis was thus in part to show the advantage of using such theory and methodology in design research, thereby improving visual design.

Paper 1:

Improving Deductive Studies of Human Behaviour in Design Through Cognitive Science – A Review and Look Forward. (Working Paper), by Andersen, E., Piccolo, S., Maier, A. To be submitted to Design Science.

Developing Deductive Studies of Human Behaviour in Design through Cognitive Science – A Review and Look Forward

Emil Andersen, Sebastiano Piccolo, Anja Maier

Abstract

The study of Human Behaviour in Design is a central part of Design Research, which has yielded a rich portfolio of insights on both the creators of artefacts and how artefacts are perceived. Recently, the design research community has seen a push towards increased rigor and focus on theory building, as a way of further developing the scientific rigor of the field. However, research is needed to determine the extent to which this ambition is needed in Human Behaviour in Design, and, if so, where and how effort could most fruitfully be exerted. Therefore, this paper investigated the topics and methods of studies on Human Behaviour in Design as a measure of the scientific rigor of the field, and as a way of suggesting improvements where needed. This was done through a comprehensive review of all deductive studies on human behaviour in seven prominent journals on design research in the time period 1979 to 2017. A total of 150 studies fulfilled the inclusion criteria and were reviewed with regards to the topics and methods used. It was found that the field was highly diverse with regards to both topics and methodology, which could slow the development of the field. Based on the characteristics of the findings, it is proposed that research could be develop more rapidly by increasingly drawing inspiration from and incorporating findings and methodology from cognitive science.

Introduction

The study of Human Behaviour in Design remains a central topic in the field of Design research, either as a topic in itself or as a complement to other areas. Overall, Human Behaviour in Design encompasses a wide variety of research ranging from designer behaviour and to behaviour while interacting with designed artefacts. In each angle, numerous contributions have been raised to our understanding of both the production and appraisal of designs. However, concerns have recently increasingly been raised with regards to how well findings within the field form a coherent picture based on theory, as well as the methodological diversity that characterises the field. Notably, the foundational editorial of this very journal saw researchers from all aspects of design research call for increased scientific rigor through stronger relation to theory and more rigorous and consistent use of methods (Papalambros et al., 2015). Their concerns have been echoed by others (Vasconceles & Crilly, 2016; Dinar et al., 2015), the most of recent of which, after thorough analyses of the field of design, warned that Design Research had two options: move towards a strong scientific theoretical foundation or risk obsolescence (Cash, 2018). As a notable example of the necessity of adaptation, Cash (2018) points to the fact that design research is seldom cited in other fields, despite the transdisciplinary nature of design research.

The present review was created to investigate the extent scientific development of Human Behaviour in Design, both with regards to theory and methodology, to identify the extent to which development is needed and to identify possible causes that may have hindered progress. Based on this, suggestions are made for future progress. While the scientific method includes both generation of ideas, building theories, formulating hypotheses and testing them (Popper, 1962; Blessing & Chakrabarti, 2009), each step being as important as the others, the present review focuses on deductive studies. This focus on the last step of the scientific process is chosen to give an overview of what is empirically tested, thus giving an overview of a solid basis upon which future studies can be built. To do so, all deductive studies in seven leading design journals were reviewed with regards to their topics and methods. We found that the field of Human Behaviour in Design has produced rich and diverse insights on designer characteristics, facilitation and obstruction of the design process and the perception of designed artefacts. Furthermore, we found that studies were highly varied in both topics and methods: Few topics and sub-topics were the subject of more than one study, and, with the exception of two studies, all studies employed newly developed paradigms. These findings indicate that theory development stands to benefit from additional studies that elaborate our understanding of existing topics. Furthermore, the findings indicate that the progress could be accelerated by employing uniform study methods that allow for easier comparison of results across studies. As a guide for this journey, we propose that researchers in Human Behaviour in Design should more extensively incorporate methods and theories from the related field of cognitive psychology. By taking as its starting point the designer as a human being, and building on the theories and methods already available in cognitive psychology, Human Behaviour in Design research is positioned to become a leading authority on both the creators and perceivers of the artificial.

Review

Review Method

The publications included in this review were selected in a three-stage process: First, titles and abstracts were read for all publications published in seven journals between their foundation and the end of 2017. The seven included journals were: International Journal of Design, Design Studies, Co-Design, Journal of Mechanical Design, Journal of Engineering Design and Research in Engineering Design and Design Science. These journals were chosen due to their prominence in the design research community. Publications were included for further review if their abstracts or titles indicated that the publications included a study of human participants and employed a deductive design. In the case of ambiguity about the study's subject material based on the abstract and title, the publication was also included for further review. Common reasons for this type of inclusion were: a) Use of subject specific subject-related terms or abbreviations that the author was not familiar with, b) Use of mathematical language (e.g. simulation) that can be used to describe human behaviour studies as well (e.g. simulation of an emergency context), and c) If a model, hypotheses, assumptions etc. were presented and the abstract did not make it clear whether the paper also included a test of these or simply discussed them. In total, 532 publications were included in the first round based on these criteria.

Second, all publications were skimmed from beginning to end by the first author to see if they applied a deductive study of human behaviour. Experiments were considered deductive if the authors at some point during the publication made statements about expected outcomes, even if these were not apparent at the beginning of the publication and/or not formulated as formal hypotheses. Notably, papers that presented a new design method were only included if the study also included predictions on human behaviour. This selection yielded a total of 150 publications: 9

from Co-Design, 3 from Design Science, 55 from Design Studies, 15 from the International Journal of Design, 15 from the of Journal of Engineering Design, 36 from the Journal of Mechanical Design, and 17 from Research in Engineering. The sample was cross-validated by the second and third author, who went through each publication to confirm that they fitted the inclusion criteria. Third, all articles were reviewed on two dimensions: topic diversity and method diversity. An overview of the review is given in Table 1.

Table 1: An overview of the papers included in each round.

Journal	1st Round	2nd Round	Percentage
Co-Design	86	10	11,63%
Design Science	11	3	27,27%
Design studies	136	55	40,44%
International Journal of Design	19	15	78,95%
Journal of Engineering Design	101	15	14,85%
Journal of Mechanical Design	72	36	50,00%
Research in Engineering Design	108	17	15,74%
Total	533	151	28,33%

Topic Portfolio

To evaluate topics researched, the studies were first categorised with regards to the subject of their study. This process revealed that the studies could be split into three categories: studies of designer characteristics, studies of obstruction and facilitation of the design process, and studies of how designed artefacts are perceived. From there, studies were further grouped based on their topic, and then further grouped into sub-topics pertaining to the specific aspects of the overarching topic. As an example of this sub-division, a study may be categorised as being a study of obstruction and facilitation of the design process, have as its topic how examples and information influence the design process, and have as its sub-topic how examples may cause design fixation. The sub-divisions were created based on the hypotheses tested in the studies and as such, papers could be represented multiple times in the review, sometimes in different categories, if they investigated several hypotheses. An overview of the studies included in the various categories, including their topics, sub-topics and experiment task are found in Table 2 for studies of Designer Characteristics, Table 3 for studies of Obstruction and Facilitation of the Design Process, and Table 4 for studies of the Perception of Designed Artefacts.

The division of studies into categories, topics and sub-topics served as the basis for evaluating the diversity of the studies: A highly uniform field would be characterised by having several studies on a few topics and sub-topics, whereas a highly diverse field would have many different topics each covered by individual or few publications.

The analyses of the diversity on the topic and sub-topic levels are detailed for each category below. As an overall finding, the vast majority of topics and sub-topics were shown to not have been the target of several studies or replication attempts. In fact, the majority of sub-topics were only the topic of a single deductive study (or only studies by the same author). While this is expected to some extent for any scientific field given that only approximately 50% of papers published get cited (Van Noorden et al., 2014), the extent of this individuality suggests a highly field overall, as will be elaborated below.

Studies of Designer Characteristics

The first category was studies of designer characteristics. These studies included in this category tested hypotheses about cognitive characteristics and strategies of designers, as well as design training. The studies predominantly involve showing how designers differ either amongst themselves or from novice designers and non-designers. A total of 45(30% of the total sample) studies were included in the category, which contained 46 hypotheses about designer characteristic spanning 8 topics and 35 sub-topics.

The topic diversity in this category was low, with 33 of the 46 (71.7%) hypotheses pertaining to one of two topics: Expert vs Novice or Designer Cognition Predicting Performance. The vast majority of papers thus either included predictions about how expert and novice designers differ from each other, or how certain traits and/or cognitive characteristics of designers are predictive of their performance.

Conversely, the sub-topic diversity was large, with only two sub-topics being investigated with more than three hypotheses, and a total of nine sub-topics being investigated with more than one hypothesis. All minor topics investigated individual sub-topics. Of the two major sub-topics, “Cognition and Performance” was the most diverse on the sub-topic level, containing only two sub-topics that had been investigated by more than one hypothesis: ‘Spatial Ability’ and ‘Ownership’. The other major topic, “Expert vs Novice” contained five sub-topics that had been investigated by more than one hypothesis, of which one had been investigated by three hypotheses, covering eleven of the nine-teen total hypotheses for the topic. This sub-topics within this topic were thus also diverse, but with several sub-topics having been the target of more than one study.

In sum, the Designer Characteristics category, while being very coherent at the topic level, consisted of hypotheses on sub-topics of which the large majority had been investigated only once, but with a trend toward having several sub-topics being the subject of several studies for the "Expert vs Novice" topic.

Studies of Obstruction and Facilitation of the Design Process

The second category was studies of methods and tools that facilitate or obstruct the design process. These studies test hypotheses about physical or cognitive techniques, stimuli and/or tools and their influence (good or bad) on the design process. The studies predominantly involve on or more designers (or design students) engaged in a design task, which was facilitated or obstructed by an experimenter induced intervention such as a text example or cognitive technique. A total of 72 (48% of the total sample) papers were included in the category, which contained 82 hypotheses about obstruction and facilitation of the design process, spanning spanned nine-teen topics and 58 sub-topics. Overall, the topic diversity in this category was high, with 40 of the 80 (50%) hypotheses pertaining to one of two topics: "Examples & Information" and "Collaboration", of which former was more than twice as frequent as the latter.

Similarly, the sub-topic diversity was large for the majority of topics. However, for the largest topic, "Examples & Information", two major sub-topics were investigated by several hypotheses: 'Fixation', which was the sub-topic of seven hypotheses, 'Distance', which was the sub-topic of four hypotheses. Furthermore, the "Design Process Characteristics" topic had five of six hypothesis related to the 'Stages' sub-topic, although this was in part due to the hypotheses being from only three papers. With these notable exceptions, few sub-topics were investigated with more than one hypothesis. As such, the Obstruction and Facilitation of Design Process category showed large topic diversity at both the topic and sub-topic level, barring some exceptions in the sub-topics 'Fixation', 'Distance' and 'Stages', which were very well investigated. Building out other sub-topics to the level of these other sub-topics could thus serve as a guiding star for future researchers. As will be discussed below, however, variations in the methods employed still could lead to a difficulty of converging on definitive statements about these well-established sub-topics.

Studies of Perception of Designed Artefacts

The third category was studies of the perception of designed artefacts. These studies test hypotheses about how participants of varying background relate to products, or about product characteristics that are relevant for design. The studies predominantly involved a product rating task with the context and artefact of choice also varying across studies. A total of 43 papers (29% of the total sample) were included in the category, which contained 55 hypotheses about Perception of

Designed Artefacts, spanning 16 topics and 48 sub-topics. Overall, the topic diversity in this category was high, with only 30 of the 52 (54.5%) hypotheses pertaining to one of three topics: "Aesthetics", "Emotion", and "Sustainability Judgments". On the sub-topic level, most topics were highly diverse, with only two sub-topics, 'Gestalt' and 'Novelty', being investigated more than twice, and an additional two sub-topics, 'Architecture Videos' and 'Categorization', being investigated more than once. Notably, within the "Aesthetics" topic, 4 of the 13 hypotheses were on the sub-topic 'Gestalt', marking at as the most well investigated sub-topic, although this was in part due to three hypotheses from the same authors. As such, the Perception of Designed Artefacts category showed large diversity at both the topic and sub-topic level, with few hypotheses relating to each topic and sub-topic as compared to the other categories, of which fewer sub-topics had been investigated by several research groups.

Looking Forward: Topics

In the following sections, we address the problems outlined above and offer suggestions on how topic coherence could be increased through integration with cognitive theory.

Increasing Replication and Elaborating Sub-Topics

The uniting finding across all categories (baring a few exceptions) was a large diversity at the sub-topic level. This diversity was furthermore also present at the topic level for the categories of Obstruction and Facilitation of the Design Process and the Perception of Designed Artefacts. This finding is notable for several reasons.

First, the low amount of hypotheses for each sub-topic is indicative of a low amount of replication of previous findings within the field. As has been seen in other fields, such as medicine and psychology, the failure to replicate findings may lead to theory building that stands on a shaky foundation (Ioannidis, 2005; Wilkinson, 1999).

Second, the low amount of hypotheses for each sub-topic suggest a less complete understanding of the various aspects of the overall topics, which could hinder the development of theories that explain the phenomena at both the overall topic level and the more narrow sub-topic level. The absence of theory building was for example reflected in the fact that no studies made inferences about the underlying mechanisms for the sub-topic phenomena. Such inferences could be highly desirable, as they would allow for spotting similarities within topics across the various sub-topics. For example, if the 'Fixation' sub-topic, which represented the most widely investigated sub-topic, a reader of a paper on 'Fixation' might learn something about how "Examples & Information" were processed in general if the findings were related to a theory of the underlying cognitive mechanism. This could in turn lead to further studies that would inform 'Fixation' through relation to the same mechanism.

Looking forward, additional studies are thus needed at the sub-topic level to confirm previous findings to solidify the foundation for future theories. This may in turn allow for inferences about the underlying mechanisms, which could unite studies on the topic and, eventually, maybe even at the category level.

Building Theories Based on Underlying Cognitive Mechanisms

The papers included in this review share in common that they propose and test hypotheses about how humans act as designers, during the design process, or with designed artefacts. Underlying all these aspects are the cognitive mechanisms that govern how we as humans experience the world, such as the attention, memory, learning and decision-making systems. The field of cognitive psychology has over the years developed elaborate theories for these cognitive mechanisms, with some mechanisms being well understood and others being heavily researched to uncover their structures. Given foundational properties of these mechanisms for human behaviour, this research represents a valuable avenue for structuring and elaborating our understanding of designers, and/or how people interact in the design process and with designed artefacts. In the following, a few key examples are listed that could be explored as an avenue for building theory based on the underlying cognitive mechanisms in the three identified categories of Human Behaviour in Design.

For studies of Designer Characteristics, several findings could be elaborated through relating them to the underlying cognitive mechanisms. For example, researchers could investigate whether the observed differences between expert and novice designers are caused by better information processing through differences in working memory, which is the cognitive mechanism that underlies conscious manipulation of information (Baddeley, 1992; 2000; 2012; Baars & Franklin, 2007; Dehaene et al., 2014). Alternatively, researcher could investigate whether the differences in how designers draw upon experience by relating them to theories of memory, such as distinguishing between implicit and explicit memory. For example, do designers draw on their experience through explicit deliberation, or does it happen automatically and implicitly more akin to remembering how to ride a bike (Ashcraft & Radvansky, 2005)?

For studying Obstruction and Facilitation of the Design Process, researchers could investigate the cognitive process that govern the observed interventions in the designers, rather than only relating the findings to the objects that facilitate or obstruct. For example, do examples stimulate memory processing or creative problem solving?

For studies of Perception of Designed Artefacts, researchers could elaborate how the various product appraisals are related to the product characteristics' priority in the attention and memory system. For example, researchers could ask whether product characteristics that are associated with higher aesthetic appeal result in easier storage in memory or higher capture of attention, as has been shown for studies of emotional stimuli and stimuli with a salient colour amongst others (Theeuwes, 1992, Öhman et al, 2001a; 2001b; Andersen & Maier, 2016; Wolfe & Horowitz, 2017).

Finally, as noted above, an advantage of using cognitive mechanisms as a framework is that it allows researchers to more easily compare across studies. In time, the use of cognitive mechanisms may thus give a holistic picture of the designer and how certain aspects of designed artefacts interact with the mechanisms underlying our general experience of the world.

Distinguishing between Conscious and Non-Conscious Behaviour

A related framework of cognitive processing lies in distinguishing between conscious and non-conscious processing and noting their relative impacts on behaviour. In particular, Dual Process theories of reasoning (Evans, 2003; 2011) and decision-making (Kahneman & Frederick, 2003) may aid design researchers in classifying the observed behaviours. While several varying theories exist under the Dual Process framework, all share in common the notion that our minds consist of two distinct systems: System 1, which governs fast, automatic and non-conscious processes, and System 2, which governs slow and deliberate processes that require conscious control (Evans, 2011). Research in cognitive psychology and related fields have used the various Dual Process frameworks to great effect in uncovering several key aspects of decision-making, such as our deviations from rationality (Tversky & Kahneman, 1974; 1981) and the possible extent and boundaries of expertise (Kahneman & Klein, 2009; Shanteau, 1988; 1992), several of which can be translated to or inspire future studies in design research.

For example, studies of designer characteristics could use Dual Process accounts to explain the transition from novice to expert, as previous research has shown that expertise is in part due to the transfer of behaviours from conscious to non-conscious control (Simon, 1992). Alternatively, studies of obstruction and facilitation of the design process could investigate whether the observed effects were due to effects on conscious or non-conscious behaviour. For example, studies could investigate whether the use of examples causes activation of System 2, thereby causing more deliberation, or whether biases such as design fixation are due to non-conscious processes that are outside conscious awareness of the designers. Studies of the perception of designs may take the distinction a step further, as the study of conscious versus non-conscious contributions to perception has been elaborated further yet within the field of consciousness research. By combining both cognitive science and neuroscience, research in this field has resulted in several theories of general consciousness (Baars & Franklin, 2007; Dehaene et al., 2014), as well as compelling evidence for the presence of several distinct types of consciousness (Block, 1995) and high level processing outside of consciousness (Hassin, 2013). As such, research on the perception of designed artefacts would benefit from knowing the relative impacts on product perception by both those characteristics that participants have report and those that lie outside of their conscious access.

Table 2: Studies of Designer Characteristics

Authors	Topic	Sub-Topic	Experiment Task(s)
Rijn et al., 2006	Cognition and Performance	Cultural Differences	Context-Mapping Task
Greenway, 1990	Cognition and Performance	Design Aptitude	Six Potential Aptitude Tests*
Yang & Cham, 2007	Cognition and Performance	Drawing Skill	Sketching Task
Johnson et al., 2014	Cognition and Performance	Empathy	Concept Generation Task
Toh et al., 2016	Cognition and Performance	Ownership	Design Task, Concept Evaluation Task
Toh et al., 2016	Cognition and Performance	Ownership	Design Task, Concept Evaluation Task
Toh & Miller, 2016	Cognition and Performance	Personality Traits	Idea Generation Task
Toh & Miller, 2016	Cognition and Performance	Risk Affinity	Idea Generation Task
Hirschi & Frey, 2002	Cognition and Performance	Short Term Memory	Problem-Solving Task
Field, 2007	Cognition and Performance	Spatial Ability	Design Course
Ho et al., 2006	Cognition and Performance	Spatial Ability	Problem-Solving Task
Ulusoy, 1999	Cognition and Performance	Spatial Ability	Design Task
Lera, 1981	Cognition and Performance	Value Theory	Design Task
Eckert et al., 2011	Designer Analysis Strategy	Notion Function	Functional Analysis Task
Nikander et al., 2014	Designer Analysis Strategy	Ownership Effect	Design Task, Design Rating
Kokotovich & Purcell, 2000	Designer vs Non-Designer	Creativity	Design Task
Powell & Newland, 1994	Designer vs Non-Designer	Learning Styles	Learning Task
Portillo & Dohr, 1989	Expert vs Novice	Advancement of Thoughts	Measure of Designing Questionnaire
Ahmed & Christensen, 2009	Expert vs Novice	Analogy Use	Design Task
Ball et al., 2004	Expert vs Novice	Analogy Use	Idea Generation Task
Gonçalves et al., 2013	Expert vs Novice	Analogy Use	Preference Assesment Task
Ahmed & Wallace, 2004	Expert vs Novice	Awareness	Design Task
Atman et al., 1999	Expert vs Novice	Behavioural Strategies	Design Task
Göker, 1997	Expert vs Novice	Behavioural Strategies	Design Task
Gosnell & Miller, 2016	Expert vs Novice	Creativity	Design Concept Evaluation Task
Menezes & Lawson, 2006	Expert vs Novice	Description	Sketch Description Task
Ahmed et al., 2003	Expert vs Novice	Design Task Approach	Design Task
Günther & Ehrlenspiel, 1999	Expert vs Novice	Education Influence	Design Task
Vermillion et al., 2015	Expert vs Novice	Education Influence	Problem-Solving Task
Viswanathan & Linsey, 2013	Expert vs Novice	Effect of Fixation	Idea Generation Task
Onarheim & Christensen, 2012	Expert vs Novice	Effect of Visual Complexity	Idea Generation Task
Bailey, 2007	Expert vs Novice	Knowledge about Design	Design Task
Verma, 1997	Expert vs Novice	Knowledge about Design	Design Course
Booth et al., 2015	Expert vs Novice	Performance	Product Dissection Task

Kim et al., 2007	Expert vs Novice	Problem Solving	Design Task
Seitamaa-Hakkarainen & Hakkarainen, 2001	Expert vs Novice	Problem Solving	Design Task
Demirbaş & Demirkan, 2003	Learning Strategy	Design Studio Stages	Design Course
Newland et al., 1987	Learning Strategy	Kolb & Leary's Theories	Kolb's LCI and LaForge & Suczek's Interpersonal Checklist
Cash & Hicks, 2013	Practice vs Lab	Performance	Design Task
Carroll et al., 1979	Problem-Solving Strategy	Decomposition	Design Task
Yilmaz et al., 2015	Problem-Solving Strategy	Heuristics	Design Task
Eckersley, 1988	Problem-Solving Strategy	Information processing	Design Task
Akin & Akin, 1996	Problem-Solving Strategy	Sudden Mental Insight	Think Aloud Design Task
Sosa & Marle, 2013	Team Characteristics & Performance	Familiarity and Creativity	Design Exercise
Cash et al., 2017	Team Characteristics & Performance	Question-asking training and Heterogenity	Design Task

*A visual memory task, a Design Task, a drawing task, a picture completion task, the Khatena-Torrance Creative Perception Inventory, the California Psychological Inventory

Table 3: Studies of Facilitation and Obstruction of the Design Process

Authors	Topic	Sub-Topic	Experiment Task(s)
Dahl et al., 2000	Cognitive Technique	Visualisation	Design task
Dahl et al., 2000	Cognitive Technique	Visualisation	Design Task
Busseri & Palmer, 2000	Collaboration	Assessment	Design Task
D'Astous et al., 2004	Collaboration	Design Evaluation Meetings	Design Evaluation Meetings
Perttula et al., 2006	Collaboration	Idea-exchanging and Performance	Idea Generation Task
Wetmore et al., 2010	Collaboration	Information Sharing	Design Review Task
Steen et al., 2013	Collaboration	Interactive Simulation	Design Task
Mitchell et al. 2016	Collaboration	Originality of Ideas	Idea Generation Task
Mitchell et al. 2016	Collaboration	Quantity of Ideas	Idea Generation Task
Cash et al., 2017	Collaboration	Question-Asking Training and Understanding	Design Task
Jang & Schunn, 2012	Collaboration	Supportive Tools	Design Task
Chulvi et al., 2017	Collaboration	Virtual vs. Face-to-Face	Design Task
Ostergaard et al., 2005	Collaboration	Virtual vs. Face-to-Face	Design Review Task
Reid & Reed, 2007	Collaboration	Visual Access	Design Task
Hanna & Barber, 2001	Computer Use	Attitude	Design Task
Savage et al., 1998	Constraints	External and Task Inherent Constraints	Design Task
Feng et al., 2014	Decision Bias	Abstraction	Phenomena Selection and Evaluation Task
Feng et al., 2014	Decision Bias	Abstraction	Phenomena Selection and Evaluation Task
Onarheim & Christensen, 2012	Decision Bias	Ownership	Idea Generation Task
Onarheim & Christensen, 2012	Decision Bias	Visual Complexity	Idea Generation Task
Dong et al., 2004	Design Documentation	Performance	Design Task
Howard et al., 2010	Design Process	Characteristics of Stages	Brainstorm Meeting
Howard et al., 2010	Design Process	Characteristics of Stages	Brainstorm Meeting
Mc Neill et al., 1998	Design Process	Characteristics of Stages	Design Task
Mc Neill et al., 1998	Design Process	Characteristics of Stages	Design Task
Yang, 2009	Design Process	Characteristics of Stages	Design Task
Alexiou et al., 2009	Design Process	Comparison to Problem-Solving	Design Task, Problem-Solving Task
Thimmaiah et al., 2017	Design Review	Design Errors	Design Review Task
Savoie & Frey, 2012	Design Review	Design of Experiments Types	Design Task
Heylighen et al., 1999	Design Studio	Conceptual Design	Design Task
Vasconcelos et al., 2017	Examples & Information	Abstract or Concrete	Idea Generation Task
Cila et al., 2014	Examples & Information	Choice of Example	Design Task
Cila et al., 2014	Examples & Information	Choice of Example	Design Task
Liikanen & Perttula, 2010	Examples & Information	Contextual Information	Idea Generation Task

Christiaans & van Andel, 1992	Examples & Information	Detail Level	Design Task
Fu et al., 2010	Examples & Information	Detail Level	Design Task
Keshwani & Chakrabarti, 2017	Examples & Information	Detail Level	Design Task
Chan et al., 2014	Examples & Information	Distance	n/a
Chiu & Shu, 2012	Examples & Information	Distance	Design Task
Keshwani & Chakrabarti, 2017	Examples & Information	Distance	Design Task
Perttula & Sipilä, 2007	Examples & Information	Distance	Design Task
Da Silva et al., 2015	Examples & Information	Examples for Ecodesigners	Design Simulation Task
Davies Cooper & Cooper, 1984	Examples & Information	Examples for Information Designers	Design Task
Atilola et al., 2015	Examples & Information	Fixation	Design Task
Jannson & Smith, 1991	Examples & Information	Fixation	Design Task
Linsey et al., 2010	Examples & Information	Fixation	Design Task
Neroni et al., 2017	Examples & Information	Fixation	Design Task
Viswanathan & Linsey, 2012	Examples & Information	Fixation	Idea Generation Task
Viswanathan & Linsey, 2013	Examples & Information	Fixation	Design Task
Viswanathan et al., 2014	Examples & Information	Fixation	Design Task
López-Mesa et al., 2011	Examples & Information	Form of Examples	Idea Generation Task
Sancar, 1996	Examples & Information	Integrative/Generative Strategy Examples	Design Task
Egan et al., 2015	Examples & Information	Inter-Level Causal Mechanisms	Design Task
Casakin & Goldschmidt, 1999	Examples & Information	Novices vs. Experts	Design Task
Egan et al., 2015	Examples & Information	Parametric Systems Relationships	Design Task
van Rijn et al., 2011	Examples & Information	Presentation Form	Design Task
López-Mesa et al., 2011	Examples & Information	Stimuli vs. Problem-Solving Style	Idea Generation Task
Atman & Bursic, 1996	Examples & Information	Teaching Materials	Design Task
Chusilp & Jin, 2005	Iteration	Conceptual Design	Design Task
Bairaktarova et al., 2017	Physical Components	Effect of Experience	Design Task
Viswanathan & Linsey, 2012	Physical Components	Mental Models	Idea Generation Task
Bairaktarova et al., 2017	Physical Components	Performance	Design Task
She & MacDonald, 2014	Priming	Ideation	Idea Generation Task
Fricke, 1999	Problem Formulation	Completeness	Design Task
Nadler et al., 1989	Problem Formulation	Coversion of Problem	Design Task
Booth et al., 2015	Product Dissection	Enumeration	Product Dissection Task
Toh et al., 2014	Product Dissection	Idea Amount and Novelty	Design Task
Toh et al., 2015	Product Dissection	Learning	Product Dissection Task
Toh et al., 2014	Product Dissection	Participation and Novelty	Design Task
Toh et al., 2014	Product Dissection	Product Type and Novelty	Design Task

Toh et al., 2015	Product Dissection	Virtual vs. Physical and Self-Efficacy	Product Dissection Task
Römer et al., 2000	Sketching	Effect on Mental Representation	Design Task
Verstijnen et al., 1998	Sketching	Effect on Process	Mental Imagery Task
Bilda & Gero, 2007	Sketching	Performance	Design Task
Schütze et al., 2003	Sketching	Performance	Design Task
Yang, 2009	Sketching	Performance	Design Task
Schaffhausen & Kowalewski, 2015	User Needs	Need Quantity pr. Person and Quality	Need Statement Task
Schaffhausen & Kowalewski, 2015	User Needs	Participant Quantity and Quality	Need Statement Task
Schaffhausen & Kowalewski, 2015	User Needs	Self-Rated Expertise and Quality	Need Statement Task
Davies, 1995	Verbalisation	Design Process	Design Task

Table 4: Perception of Designed Artefacts

Authors	Topic	Sub-Topic	Experiment Task(s)
Stamps,1999	Aesthetics	Architechture Detail Perception and Van der Laan Septives	Pair-Wise Evaluation Task
Al-Azzawi et al., 2007	Aesthetics	Comparison of Options	Free Sorting Task
Blijlevens et al., 2013	Aesthetics	Deviation from 'Prototype'	Product Evaluation Task
Lugo et al., 2016	Aesthetics	Gestalt	Product Evaluation Task
Valencia-Romero & Lugo, 2017	Aesthetics	Gestalt	Product Evaluation Task
Valencia-Romero & Lugo, 2017	Aesthetics	Gestalt	Product Evaluation Task
Valencia-Romero & Lugo, 2017	Aesthetics	Gestalt	Product Evaluation Task
Faerbar & Carbon, 2013	Aesthetics	Imitation vs. Original	Product Evaluation Task
Da Silva et al., 2015	Aesthetics	Knowledge of Designer Intention	Product Evaluation Task
Coughlan & Mashman, 1999	Aesthetics	Novelty	Design Evaluation Task
Hung & Chen, 2012	Aesthetics	Novelty	Sorting Task
Ludden et al., 2012	Aesthetics	Visual-Tactual Incongruity	Product Interaction
Phillips, 1982	Child Perception	Design Education	Product Evaluation Task
Phillips, 1982	Child Perception	Gender	Product Evaluation Task
Du & MacDonald, 2014	Cognition Predicting Perception	Importance Rating	Feature Rating Task
Du & MacDonald, 2014	Cognition Predicting Perception	Noticibility of Change	Feature Rating Task
Dong & Lee, 2008	Cognition Predicting Perception	Webpage Usage	Webpage Use Task
Bezawada et al., 2017	Comfort	Physical Equipment	Engineering Task
Blijlevens et al., 2009	Distinguishing Between Products	Categorization	Product Evaluation Task
Jordan & Persson, 2007	Distinguishing Between Products	Categorization	Odd-One-Out Selection Task
Demir et al., 2009	Emotion	Appraisal Theory	Product Interaction Task
Spence & Zampini, 2007	Emotion	Auditory Cues	Product Evaluation Task
Miesler et al., 2011	Emotion	Cuteness	Product Evaluation Task
Mata et al., 2017	Emotion	Desire to Own	Product Evaluation Task
Mugge et al., 2009	Emotion	Effort	Product Evaluation Task
Desmet, 2012	Emotion	Positive Emotion	Product Evaluation Task
Munoz & Tucker, 2016	Emotion	Semantic Structure	Lecture Viewing
Mata et al., 2017	Emotion	Shape	Product Evaluation Task
Tsai et al., 2008	Emotion	Uncertainty	Product Evaluation Task
Caldwell et al., 2012	Interpretability	Function Type	Classification Task
Caldwell et al., 2012	Interpretability	Language Specificity	Classification Task
Blackler et al., 2003	Intuitiveness	Experience with Product	Product Use Task
Lin & Yang, 2010	Metaphor	Memory	Product Evaluation Task

Rhodes & Powell, 1994	Narrative	Architecture Videos	Video Retelling Task
Rhodes et al., 1992	Narrative	Architecture Videos	Video Retelling Task
Wei et al.,2014	Perceived Quality	Colour	Product Evaluation Task
Mugge & Schoormans, 2012	Perceived Quality	Novelty	Product Evaluation Task
Jordan & Persson, 2007	Perception style	Product Types	Product Evaluation Task
Reid et al., 2013	Presentation Medium	Computer Sketches, Simplified Renderings and Silhouettes	Product Evaluation Task
Du & MacDonald, 2015	Purchase Intention	Cancellation-and-Focus Decision Model	Product Evaluation Task
Wan, 1987	Purchase Intention	Cognitive Dissonance	Product Evaluation Task
MacDonald et al., 2009	Purchase Intention	Crux and Sentinel Attributes	Product Evaluation Task
Cho & Kim, 2012	Purchase Intention	Flow	Website User Test
Cho & Kim, 2012	Purchase Intention	Self-Congrueity	Website User Test
Lee et al., 2015	Sustainability Judgments	Car Colour	Product Evaluation Task
Lee et al., 2015	Sustainability Judgments	Car Emblem	Product Evaluation Task
Goucher-Lambert et al., 2017	Sustainability Judgments	Emotion and Moral	Product Evaluation Task
Goucher-Lambert & Cagan, 2015	Sustainability Judgments	Knowledge of Impact	Product Evaluation Task
Reid et al., 2010	Sustainability Judgments	Line Changes	Product Evaluation Task
Goucher-Lambert & Cagan, 2015	Sustainability Judgments	Monetary Investment	Product Evaluation Task
Reid et al., 2010	Sustainability Judgments	Nature Likeness	Product Evaluation Task
Sacharin et al., 2011	Sustainability Judgments	Safety	Product Evaluation Task
Jakesch et al., 2011	Tactile evalation	Multi-Modality & Performance	Sensory Evaluation Task

Method Portfolio

As our second measure of diversity, we investigated the degree to which studies employed the same or similar methods to each other. This measure was included as lower diversity in methods vastly simplifies comparison across studies, thereby allowing for easier interpretation on the strength of evidence on the relationship between phenomena. While having several studies on the same topic and sub-topic is important for elucidating the various aspects of the constructs and for ensuring that findings are not due to random observations, as elaborated in the above discussion on replication, it is thus similarly important that the studies share methodology that allows for the studies to be comparable. Therefore, the diversity of methods, meaning sampling methods, experiment tasks, and outcome variables were reviewed.

Sampling methods

To measure the diversity of sampling methods, we classified each study with regards to the sampling method, the experiment task that the participants would be performing, and the outcome variable analysed to test the hypothesis. Furthermore, the exact methods employed were noted down and investigated to evaluate the comparability across studies. Of the 150 studies included in the review, two studies used the exact same sampling method (Kolb's LCI) to measure their outcome (Newland et al., 1987; Powell & Newland, 1994). As such, while the overall method category and object of the study was fairly consistent across categories, all but two studies created their own paradigm from scratch to investigate their hypothesis/hypotheses. Three major problems could arise with such diversity of outcomes:

First, for studies involving design tasks, wherein participants (usually designers or design students) complete a design assignment, the discrepancy across studies in the exact design task employed is problematic due to the variable nature of design. Indeed, a considerable body of work has described how different design tasks are associated with different ability requirements, cognitive employment amongst others. However, the fact that newly developed design tasks were employed in all design tasks in the reviewed studies suggests that design tasks are assumed to be sufficiently homogeneous to be compared, even if they differ substantially. Given such a discrepancy, it seems worthwhile for future research to clarify the extent to which design tasks can differ without significant adverse effects to the comparability of their outcomes. In a previous review on Design Fixation (the 'Fixation' sub-topic in this review), Vasconcelos & Crilly (2016) noted that the phenomena had been investigated with such diverse methods that results were barely comparable across studies. They concluded that this may have caused the prevalence of contradictory findings with regards to the effect of design fixation, and that these contradictions could not reasonably be resolved without new studies that employed more similar designs. Given that 'Fixation' is the sub-topic that has been investigated by the most hypotheses, it should be expected that other topics and sub-topics would be subject to the difficulties with regards to interpretation. In other words, if one study contradicts another, a lingering doubt should remain with regards whether it was due to differences in sampling that was the cause, as long as the study methodologies differ.

As an extension of this, the second problem with design tasks lies in the rating system employed in most of these studies. A considerable body of research has shown that designers are quite unique with regards to style and evaluations. Indeed, one may consider this a major contribution of the design research field. However, when it comes to employing "expert evaluation", which our results show to be a very common method, as the measure of the quality of participant's design, these differences are not taken into account. As with differences in design tasks, such possible differences in expert evaluations puts into question whether the findings were due to the proposed phenomena or the characteristics of the evaluator.

Third, a similar problem can be observed with regards to cognitive measures of designers. A considerable amount of effort is dedicated to establishing the minute differences across different cognitive aspects in psychology. As the most well-known example, IQ is measured differently by different IQ tests, with the WAIS measuring full-range IQ and Raven's Matrices measuring mostly mathematical spatial abilities. While both relate to the same construct (IQ), no psychologist would compare two studies using these two different tests without considerable consideration. In the same vein, some hesitation should be included before interpreting across cognitive measures that were not deduced from the same instruments.

Looking Forward Through Cognitive Psychology: Method Coherence

In the following, we offer suggestions from cognitive psychology for alleviating the above stated problems of method coherence. It is suggested that many problems could be solved rapidly by the individual researchers, while others will require larger efforts through dedicated future projects.

Replicating Previous Studies

In the same manner that replicating studies with regards to topic and sub-topic could improve understanding and reliability of previous findings, replication of methodology stands to improve understanding and reliability of the topic investigated by increasing the comparability across studies. To do so, researchers should look to replicate studies using the exact paradigms of preceding studies. When doing variations of a study, researchers should attempt to change as few things as possible and only with a specific purpose in mind. As an example of done effectively, Brosch & Sharm (2005) replicated Öhman et al., (2001), but also performed a variation of the original study where stimuli were changed while maintaining the structure. As a result, Brosch & Sharma could confidently say that the variation were due to the variation in stimuli. Another prominent example from attention research comes in Benoni & Tsal (2010), who ensured to first replicate the findings of (Lavie & Tsal, 1995), with a paradigm that closely mimicked the studies, but allowed for variations. This approach allowed the authors to make confident statements about the discrepancy between results, eventually resulting in a different theory being proposed. Similarly, studies in design research could replicate e.g. the design tasks used in previous studies, thereby creating increased confidence when analysing any discrepancies between studies that may arise.

Developing Robust Paradigms

Naturally, one may ask "which study should serve as the base for replication?". Given that several studies have been invented at the present stage of design research, this is very valid question indeed. In the short term, research may simply choose whichever study best fits their intentions. In the long term however, efforts could be made to have dedicated research projects that have as their objective to create methods that are robust and measure their constructs reliably. Such studies represent a large proportion of the psychological literature, many of which end up in commercially viable products. Design research may then be an authority on core subjects not only due to the published research, but also as a result of their reliable measures of that can be useful for other research groups.

Conclusion

Research in Human Behaviour in Design remains a vital part of the design research community. Several decades of deductive studies have been conducted, yielding substantial insights into both the creators and perceivers of the artificial. A review was conducted of all deductive studies in seven prominent journals in the period 1979-2017 to investigate the portfolio of topics and methods as a measure of scientific rigour. The results showed that, upon further inspection, the field is highly diverse with regards to topics, sub-topics and methodology. However, by building future research on theoretical and methodological approaches used in the neighbouring field of cognitive psychology, the field of Human Behaviour in Design is positioned to become a leading authority on both the creators and perceivers of the artificial.

It was found that the studies were highly diverse at the sub-topic level in all categories, and at the topic level for studies of Obstruction and Facilitation of the Design Process and Perception of Designed Artefacts, which could lead to slower theory development. For the method review, each category was reviewed on the diversity in experiment tasks, sampling methods and outcome variables. It was found that the field was highly diverse in methodology used, with only two studies employing the exact same methodology. For both the findings on topic and methods, it is proposed that research could be improved by drawing inspiration from and incorporating findings and methodology from cognitive science.

References

1. Cash, P. J. (2018). Developing theory-driven design research. *Design Studies*, *56*, 84–119. https://doi.org/10.1016/j.destud.2018.03.002
2. Dinar, M., Shah, J. J., Cagan, J., Leifer, L., Linsey, J., Smith, S. M., & Hernandez, N. V. (2015). Empirical Studies of Designer Thinking: Past, Present, and Future. *Journal of Mechanical Design*, *137*(2), 021101–021101. https://doi.org/10.1115/1.4029025
3. Ioannidis, J. P. A. (2005). Contradicted and Initially Stronger Effects in Highly Cited Clinical Research. *JAMA*, *294*(2), 218–228. https://doi.org/10.1001/jama.294.2.218
4. Katsikopoulos, K. V. (2012). Decision Methods for Design: Insights from Psychology. *Journal of Mechanical Design*, *134*(8), 084504. https://doi.org/10.1115/1.4007001
5. Melnyk, S. (1998). May you live in interesting times…the emergence of theory-driven empirical research. *Journal of Operations Management*, *16*(4), 311–319. https://doi.org/10.1016/S0272-6963(98)00027-8
6. Papalambros, P. Y. (2015). Design Science: Why, What and How. *Design Science*, *1*. https://doi.org/10.1017/dsj.2015.1
7. Van Noorden, R., Maher, B., & Nuzzo, R. (2014). The top 100 papers. *Nature News*, *514*(7524), 550. https://doi.org/10.1038/514550a
8. Wilkinson, L. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, *54*(8), 594.
9. Cash, P. J. (2018). Developing theory-driven design research. *Design Studies*, *56*, 84–119. https://doi.org/10.1016/j.destud.2018.03.002
10. Dinar, M., Shah, J. J., Cagan, J., Leifer, L., Linsey, J., Smith, S. M., & Hernandez, N. V. (2015). Empirical Studies of Designer Thinking: Past, Present, and Future. *Journal of Mechanical Design*, *137*(2), 021101–021101. https://doi.org/10.1115/1.4029025
11. Ioannidis, J. P. A. (2005). Contradicted and Initially Stronger Effects in Highly Cited Clinical Research. *JAMA*, *294*(2), 218–228. https://doi.org/10.1001/jama.294.2.218
12. Katsikopoulos, K. V. (2012). Decision Methods for Design: Insights from Psychology. *Journal of Mechanical Design*, *134*(8), 084504. https://doi.org/10.1115/1.4007001
13. Melnyk, S. (1998). May you live in interesting times…the emergence of theory-driven empirical research. *Journal of Operations Management*, *16*(4), 311–319. https://doi.org/10.1016/S0272-6963(98)00027-8

14. Papalambros, P. Y. (2015). Design Science: Why, What and How. *Design Science*, *1*. https://doi.org/10.1017/dsj.2015.1
15. Van Noorden, R., Maher, B., & Nuzzo, R. (2014). The top 100 papers. *Nature News*, *514*(7524), 550. https://doi.org/10.1038/514550a
16. Wilkinson, L. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, *54*(8), 594.
17. Ahmed, S. (2004). Understanding the knowledge needs of novice designers in the aerospace industry. *Design Studies*, *25*(2), 155–173. https://doi.org/10.1016/j.destud.2003.10.006
18. Ahmed, Saeema, & Christensen, B. T. (2009). An In Situ Study of Analogical Reasoning in Novice and Experienced Design Engineers. *Journal of Mechanical Design*, *131*(11), 111004-111004–111009. https://doi.org/10.1115/1.3184693
19. Ahmed, Saeema, Wallace, K. M., & Blessing, L. T. (2003). Understanding the differences between how novice and experienced designers approach design tasks. *Research in Engineering Design*, *14*(1), 1–11. https://doi.org/10.1007/s00163-002-0023-z
20. Akin, Ö., & Akin, C. (1996). Frames of reference in architectural design: analysing the hyperacclamation (A-h-a-!). *Design Studies*, *17*(4), 341–361. https://doi.org/10.1016/S0142-694X(96)00024-5
21. Al-Azzawi, A., Frohlich, D., & Wilson, M. (2007). Beauty constructs for MP3 players. *CoDesign*, *3*(sup1), 59–74. https://doi.org/10.1080/15710880701334447
22. Alexiou, K., Zamenopoulos, T., Johnson, J. H., & Gilbert, S. J. (2009). Exploring the neurological basis of design cognition using brain imaging: some preliminary results. *Design Studies*, *30*(6), 623–647. https://doi.org/10.1016/j.destud.2009.05.002
23. Atilola, O., Tomko, M., & Linsey, J. S. (2016). The effects of representation on idea generation and design fixation: A study comparing sketches and function trees. *Design Studies*, *42*(Supplement C), 110–136. https://doi.org/10.1016/j.destud.2015.10.005
24. Atman, C. J., & Bursic, K. M. (1996). Teaching engineering design: Can reading a textbook make a difference? *Research in Engineering Design*, *8*(4), 240–250. https://doi.org/10.1007/BF01597230
25. Atman, C. J., Chimka, J. R., Bursic, K. M., & Nachtmann, H. L. (1999). A comparison of freshman and senior engineering design processes. *Design Studies*, *20*(2), 131–152. https://doi.org/10.1016/S0142-694X(98)00031-3
26. Bailey, R. (2007). Effects of Industrial Experience and Coursework During Sophomore and Junior Years on Student Learning of Engineering Design. *Journal of Mechanical Design*, *129*(7), 662–667. https://doi.org/10.1115/1.2722323
27. Bairaktarova, D., Graziano, W., & Cox, M. (2017). Enhancing Engineering Students' Performance on Design Task: The Box of Parts. *Journal of Mechanical Design*, *139*(5), 052001-052001–052009. https://doi.org/10.1115/1.4036128
28. Ball, L. J., Ormerod, T. C., & Morley, N. J. (2004). Spontaneous analogising in engineering design: a comparative analysis of experts and novices. *Design Studies*, *25*(5), 495–508. https://doi.org/10.1016/j.destud.2004.05.004
29. Bezawada, S., Hu, Q., Gray, A., Brick, T., & Tucker, C. (2017). Automatic Facial Feature Extraction for Predicting Designers' Comfort With Engineering Equipment During Prototype Creation. *Journal of Mechanical Design*, *139*(2), 021102-021102–021110. https://doi.org/10.1115/1.4035428
30. Bilda, Z., & Gero, J. S. (2007). The impact of working memory limitations on the design process during conceptualization. *Design Studies*, *28*(4), 343–367. https://doi.org/10.1016/j.destud.2007.02.005
31. Blackler, A., Popovic, V., & Mahar, D. (2003). The nature of intuitive use of products: an experimental approach. *Design Studies*, *24*(6), 491–506. https://doi.org/10.1016/S0142-694X(03)00038-3
32. Blijlevens, J., Creusen, M. E., & Schoormans, J. P. (2009). How consumers perceive product appearance: The identification of three product appearance attributes. *International Journal of Design*, *3*(3). Retrieved from http://search.proquest.com/openview/859b4e34e104b3ac81d29cce1737f962/1?pq-origsite=gscholar
33. Blijlevens, J., Mugge, R., Ye, P., & Schoormans, J. P. (2013). The influence of product exposure on trendiness and aesthetic appraisal. *International Journal of Design*, *7*(1).
34. Booth, J. W., Reid, T. N., Eckert, C., & Ramani, K. (2015). Comparing Functional Analysis Methods for Product Dissection Tasks. *Journal of Mechanical Design*, *137*(8), 081101-081101–081110. https://doi.org/10.1115/1.4030232
35. Busseri, M. A., & Palmer, J. M. (2000). Improving teamwork: the effect of self-assessment on construction design teams. *Design Studies*, *21*(3), 223–238. https://doi.org/10.1016/S0142-694X(99)00024-1
36. Caldwell, B. W., Thomas, J. E., Sen, C., Mocko, G. M., & Summers, J. D. (2012). The Effects of Language and Pruning on Function Structure Interpretability. *Journal of Mechanical Design*, *134*(6), 061001-061001–061011. https://doi.org/10.1115/1.4006442
37. Carroll, J. M., Thomas, J. C., & Malhotra, A. (1979). Clinical— experimental analysis of design problem solving. *Design Studies*, *1*(2), 84–92. https://doi.org/10.1016/0142-694X(79)90004-8

38. Casakin, H., & Goldschmidt, G. (1999). Expertise and the use of visual analogy: implications for design education. *Design Studies*, *20*(2), 153–175. https://doi.org/10.1016/S0142-694X(98)00032-5
39. Cash, P., Dekoninck, E. A., & Ahmed-Kristensen, S. (2017). Supporting the development of shared understanding in distributed design teams. *Journal of Engineering Design*, *28*(3), 147–170. https://doi.org/10.1080/09544828.2016.1274719
40. Cash, P. J., Hicks, B. J., & Culley, S. J. (2013). A comparison of designer activity using core design situations in the laboratory and practice. *Design Studies*, *34*(5), 575–611. https://doi.org/10.1016/j.destud.2013.03.002
41. Chan, J., Dow, S. P., & Schunn, C. D. (2014). Do the best design ideas (really) come from conceptually distant sources of inspiration? *Design Studies*, *36*(Supplement C), 31–58. https://doi.org/10.1016/j.destud.2014.08.001
42. Chiu, I., & Shu, L. H. (2012). Investigating effects of oppositely related semantic stimuli on design concept creativity. *Journal of Engineering Design*, *23*(4), 271–296. https://doi.org/10.1080/09544828.2011.603298
43. Cho, E., & Kim, Youn-Kyung. (2012). The Effects of Website Designs, Self-Congruity, and Flow on Behavioral Intention. *International Journal of Design*, *6*(2), 31–39.
44. Christiaans, H., & van Andel, J. (1993). The effects of examples on the use of knowledge in a student design activity: the case of the 'flying Dutchman'. *Design Studies*, *14*(1), 58–74. https://doi.org/10.1016/S0142-694X(05)80005-5
45. Chulvi, V., Mulet, E., Felip, F., & García-García, C. (2017). The effect of information and communication technologies on creativity in collaborative design. *Research in Engineering Design*, *28*(1), 7–23. https://doi.org/10.1007/s00163-016-0227-2
46. Chusilp, P., & Jin, Y. (2005). Impact of Mental Iteration on Concept Generation. *Journal of Mechanical Design*, *128*(1), 14–25. https://doi.org/10.1115/1.2118707
47. Cila, N., Hekkert, P. P. M., & Visch, V. T. (2014). Source selection in product metaphor generation: The effect of salience and relatedness. *International Journal of Design 8 (1), 15-28.(2014)*. Retrieved from http://repository.tudelft.nl/view/ir/uuid:190d0579-fd25-476b-bcde-1c0d39219571/
48. Corremans, J. A. M. (2011). Measuring the effectiveness of a design method to generate form alternatives: an experiment performed with freshmen students product development. *Journal of Engineering Design*, *22*(4), 259–274. https://doi.org/10.1080/09544820903312416
49. Coughlan, P., & Mashman, R. (1999). Once is not enough: repeated exposure to and aesthetic evaluation of an automobile design prototype. *Design Studies*, *20*(6), 553–563. https://doi.org/10.1016/S0142-694X(99)00007-1
50. da Silva, M. K., Remy, S., & Reyes, T. (2015). On providing design process information to the environmental expert. *Research in Engineering Design*, *26*(4), 327–336. https://doi.org/10.1007/s00163-015-0198-8
51. da Silva, O., Crilly, N., & Hekkert, P. (2015). How people's appreciation of products is affected by their knowledge of the designers' intentions. *International Journal of Design*. Retrieved from https://www.repository.cam.ac.uk/handle/1810/253560
52. Dahl, D. W., Chattopadhyay, A., & Gorn, G. J. (2001). The importance of visualisation in concept design. *Design Studies*, *22*(1), 5–26. https://doi.org/10.1016/S0142-694X(99)00028-9
53. D'Astous, P., Détienne, F., Visser, W., & Robillard, P. N. (2004). Changing our view on design evaluation meetings methodology: a study of software technical review meetings. *Design Studies*, *25*(6), 625–655. https://doi.org/10.1016/j.destud.2003.12.002
54. Davies Cooper, R., & Cooper, C. L. (1984). The evaluation of guidelines to aid the information designer. *Design Studies*, *5*(4), 248–267. https://doi.org/10.1016/0142-694X(84)90060-7
55. Davies, S. P. (1995). Effects of concurrent verbalization on design problem solving. *Design Studies*, *16*(1), 102–116. https://doi.org/10.1016/0142-694X(95)90649-Z
56. Demir, E., Desmet, P. M., & Hekkert, P. (2009). Appraisal patterns of emotions in human-product interaction. *International Journal of Design*, *3*(2). Retrieved from http://search.proquest.com/openview/29d3d6e71340ff9008e64f16a0139c2c/1?pq-origsite=gscholar
57. Demirbaş, O. O., & Demirkan, H. (2003). Focus on architectural design process through learning styles. *Design Studies*, *24*(5), 437–456. https://doi.org/10.1016/S0142-694X(03)00013-9
58. Desmet, P. M. A. (2012). Faces of Product Pleasure: 25 Positive Emotions in Human-Product Interactions. *International Journal of Design*, *6*(2), 1–29.
59. Dong, A., Hill, A. W., & Agogino, A. M. (2004). A Document Analysis Method for Characterizing Design Team Performance. *Journal of Mechanical Design*, *126*(3), 378–385. https://doi.org/10.1115/1.1711818
60. Dong, Y., & Lee, K.-P. (2008). A cross-cultural comparative study of users' perceptions of a webpage: With a focus on the cognitive styles of Chinese, Koreans and Americans. *International Journal of Design*, *2*(2), 19–30.
61. Du, P., & MacDonald, E. F. (2014). Eye-Tracking Data Predict Importance of Product Features and Saliency of Size Change. *Journal of Mechanical Design*, *136*(8), 081005-081005–081014. https://doi.org/10.1115/1.4027387
62. Du, P., & MacDonald, E. F. (2015). Products' Shared Visual Features Do Not Cancel in Consumer Decisions. *Journal of Mechanical Design*, *137*(7), 071409-071409–071411. https://doi.org/10.1115/1.4030162

63. Eckersley, M. (1988). The form of design processes: a protocol analysis study. *Design Studies*, *9*(2), 86–94. https://doi.org/10.1016/0142-694X(88)90034-8
64. Eckert, C., Alink, T., Ruckpaul, A., & Albers, A. (2011). Different notions of function: results from an experiment on the analysis of an existing product. *Journal of Engineering Design*, *22*(11–12), 811–837. https://doi.org/10.1080/09544828.2011.603297
65. Egan, P., Schunn, C., Cagan, J., & LeDuc, P. (2015). Improving human understanding and design of complex multi-level systems with animation and parametric relationship supports. *Design Science*, *1*. https://doi.org/10.1017/dsj.2015.3
66. Faerber, S. J., & Carbon, C.-C. (2013). Jump on the innovator's train: cognitive principles for creating appreciation in innovative product designs. *Research in Engineering Design*, *24*(3), 313–319. https://doi.org/10.1007/s00163-012-0148-7
67. Feng, T., Cheong, H., & Shu, L. H. (2014). Effects of Abstraction on Selecting Relevant Biological Phenomena for Biomimetic Design. *Journal of Mechanical Design*, *136*(11), 111111-111111–10. https://doi.org/10.1115/1.4028173
68. Field, B. W. (2007). Visualization, Intuition, and Mathematics Metrics as Predictors of Undergraduate Engineering Design Performance. *Journal of Mechanical Design*, *129*(7), 735–743. https://doi.org/10.1115/1.2722790
69. Fricke, G. (1999). Successful approaches in dealing with differently precise design problems. *Design Studies*, *20*(5), 417–429. https://doi.org/10.1016/S0142-694X(99)00018-6
70. Fu, K., Cagan, J., & Kotovsky, K. (2010). Design Team Convergence: The Influence of Example Solution Quality. *Journal of Mechanical Design*, *132*(11), 111005-111005–111011. https://doi.org/10.1115/1.4002202
71. Göker, M. H. (1997). The effects of experience during design problem solving. *Design Studies*, *18*(4), 405–426. https://doi.org/10.1016/S0142-694X(97)00009-4
72. Gonçalves, M., Cardoso, C., & Badke-Schaub, P. (2014). What inspires designers? Preferences on inspirational approaches during idea generation. *Design Studies*, *35*(1), 29–53. https://doi.org/10.1016/j.destud.2013.09.001
73. Gosnell, C. A., & Miller, S. R. (2015). But Is It Creative? Delineating the Impact of Expertise and Concept Ratings on Creative Concept Selection. *Journal of Mechanical Design*, *138*(2), 021101-021101–021111. https://doi.org/10.1115/1.4031904
74. Goucher-Lambert, K., & Cagan, J. (2015). The Impact of Sustainability on Consumer Preference Judgments of Product Attributes. *Journal of Mechanical Design*, *137*(8), 081401-081401–081411. https://doi.org/10.1115/1.4030271
75. Goucher-Lambert, K., Moss, J., & Cagan, J. (2017). Inside the Mind: Using Neuroimaging to Understand Moral Product Preference Judgments Involving Sustainability. *Journal of Mechanical Design*, *139*(4), 041103-041103–041111. https://doi.org/10.1115/1.4035859
76. Greenway, P. (1990). Selecting landscape architecture students: measuring design aptitude. *Design Studies*, *11*(2), 89–95. https://doi.org/10.1016/0142-694X(90)90023-6
77. Gregory, S. A. (1982). Design as part of a job. *Design Studies*, *3*(2), 65–76. https://doi.org/10.1016/0142-694X(82)90051-5
78. Günther, J., & Ehrlenspiel, K. (1999). Comparing designers from practice and designers with systematic design education. *Design Studies*, *20*(5), 439–451. https://doi.org/10.1016/S0142-694X(99)00019-8
79. Hanna, R., & Barber, T. (2001). An inquiry into computers in design: attitudes before–attitudes after. *Design Studies*, *22*(3), 255–281. https://doi.org/10.1016/S0142-694X(00)00029-6
80. Heylighen, A, & Verstijnen, I. M. (2003). Close encounters of the architectural kind. *Design Studies*, *24*(4), 313–326. https://doi.org/10.1016/S0142-694X(02)00040-6
81. Heylighen, Ann, Bouwen, J. E., & Neuckermans, H. (1999). Walking on a thin line—Between passive knowledge and active knowing of components and concepts in architectural design: Between passive knowledge and active knowing of components and concepts in architectural design. *Design Studies*, *20*(2), 211–235. https://doi.org/10.1016/S0142-694X(98)00035-0
82. Hirschi, N., & Frey, D. (2002). Cognition and complexity: An experiment on the effect of coupling in parameter design. *Research in Engineering Design*, *13*(3), 123–131. https://doi.org/10.1007/s00163-002-0011-3
83. Ho, C.-H., Eastman, C., & Catrambone, R. (2006). An investigation of 2D and 3D spatial and mathematical abilities. *Design Studies*, *27*(4), 505–524. https://doi.org/10.1016/j.destud.2005.11.007
84. Howard, T. J., Dekoninck, E. A., & Culley, S. J. (2010). The use of creative stimuli at early stages of industrial product innovation. *Research in Engineering Design*, *21*(4), 263–274. https://doi.org/10.1007/s00163-010-0091-4
85. Hung, W.-K., & Chen, L.-L. (2012). Effects of Novelty and Its Dimensions on Aesthetic Preference in Product Design. *International Journal of Design*, *6*(2), 81–90.
86. Ip, B., & Jacobs, G. (2004). Quantifying game design. *Design Studies*, *25*(6), 607–624. https://doi.org/10.1016/j.destud.2004.02.001

87. Jakesch, M., Zachhuber, M., Leder, H., Spingler, M., & Carbon, C.-C. (2011). Scenario-based touching: on the influence of top-down processes on tactile and visual appreciation. *Research in Engineering Design*, *22*(3), 143–152. https://doi.org/10.1007/s00163-010-0102-5
88. Jang, J., & Schunn, C. D. (2012). Physical Design Tools Support and Hinder Innovative Engineering Design. *Journal of Mechanical Design*, *134*(4), 041001-041001–041009. https://doi.org/10.1115/1.4005651
89. Jansson, D. G., & Smith, S. M. (1991). Design fixation. *Design Studies*, *12*(1), 3–11. https://doi.org/10.1016/0142-694X(91)90003-F
90. Johnson, D. G., Genco, N., Saunders, M. N., Williams, P., Seepersad, C. C., & Hölttä-Otto, K. (2014). An Experimental Investigation of the Effectiveness of Empathic Experience Design for Innovative Concept Generation. *Journal of Mechanical Design*, *136*(5), 051009-051009–051012. https://doi.org/10.1115/1.4026951
91. Jordan, P. W., & Persson, S. (2007). Exploring users' product constructs: how people think about different types of product. *CoDesign*, *3*(sup1), 97–106. https://doi.org/10.1080/15710880701327862
92. Keshwani, S., & Chakrabarti, A. (2017). Influence of analogical domains and comprehensiveness in explanation of analogy on the novelty of designs. *Research in Engineering Design*, *28*(3), 381–410. https://doi.org/10.1007/s00163-016-0246-z
93. Kim, M. H., Kim, Y. S., Lee, H. S., & Park, J. A. (2007). An underlying cognitive aspect of design creativity: Limited Commitment Mode control strategy. *Design Studies*, *28*(6), 585–604. https://doi.org/10.1016/j.destud.2007.04.006
94. Kokotovich, V., & Purcell, T. (2000). Mental synthesis and creativity in design: an experimental examination. *Design Studies*, *21*(5), 437–449. https://doi.org/10.1016/S0142-694X(00)00017-X
95. Lee, J., Jung, B., & Chu, W. (2015). Signaling environmental altruism through design: The role of green cue prominence in hybrid cars. *International Journal of Design*, *9*(2), 79–91.
96. Lee, T. Y., & Radcliffe, D. F. (1990). Innate design abilities of first year engineering and industrial design students. *Design Studies*, *11*(2), 96–106. https://doi.org/10.1016/0142-694X(90)90024-7
97. Lera, S. G. (1981). Architectural designers' values and the evaluation of their designs. *Design Studies*, *2*(3), 131–137. https://doi.org/10.1016/0142-694X(81)90067-3
98. Liikkanen, L. A., & Perttula, M. (2010). Inspiring design idea generation: insights from a memory-search perspective. *Journal of Engineering Design*, *21*(5), 545–560. https://doi.org/10.1080/09544820802353297
99. Lin, P.-C., & Yang, Chao-Ming. (2010). Impact of Product Pictures and Brand Names on Memory of Chinese Metaphorical Advertisements. *International Journal of Design*, *4*(1), 57–70.
100. Linsey, J. S., Tseng, I., Fu, K., Cagan, J., Wood, K. L., & Schunn, C. (2010). A Study of Design Fixation, Its Mitigation and Perception in Engineering Design Faculty. *Journal of Mechanical Design*, *132*(4), 041003-041003–041012. https://doi.org/10.1115/1.4001110
101. López-Mesa, B., Mulet, E., Vidal, R., & Thompson, G. (2011). Effects of additional stimuli on idea-finding in design teams. *Journal of Engineering Design*, *22*(1), 31–54. https://doi.org/10.1080/09544820902911366
102. Ludden, G. D., Schifferstein, H. N., & Hekkert, P. (2012). Beyond surprise: A longitudinal study on the experience of visual-tactual incongruities in products. *International Journal of Design*, *6*(1). Retrieved from http://search.proquest.com/openview/3271681ef21934c19d6b2430d6cbaac7/1?pq-origsite=gscholar
103. Lugo, J. E., Schmiedeler, J. P., Batill, S. M., & Carlson, L. (2016). Relationship Between Product Aesthetic Subject Preference and Quantified Gestalt Principles in Automobile Wheel Rims. *Journal of Mechanical Design*, *138*(5), 051101-051101–051110. https://doi.org/10.1115/1.4032775
104. MacDonald, E. F., Gonzalez, R., & Papalambros, P. (2009). The construction of preferences for crux and sentinel product attributes. *Journal of Engineering Design*, *20*(6), 609–626. https://doi.org/10.1080/09544820802132428
105. Mata, M. P., Ahmed-Kristensen, S., Brockhoff, P. B., & Yanagisawa, H. (2017). Investigating the influence of product perception and geometric features. *Research in Engineering Design*, *28*(3), 357–379. https://doi.org/10.1007/s00163-016-0244-1
106. Mc Neill, T., Gero, J. S., & Warren, J. (1998). Understanding conceptual electronic design using protocol analysis. *Research in Engineering Design*, *10*(3), 129–140. https://doi.org/10.1007/BF01607155
107. Menezes, A., & Lawson, B. (2006). How designers perceive sketches. *Design Studies*, *27*(5), 571–585. https://doi.org/10.1016/j.destud.2006.02.001
108. Miesler, L. (2011). Isn't It Cute: An Evolutionary Perspective of Baby-Schema Effects in Visual Product Designs. *International Journal of Design*, *5*(3), 17–30.
109. Mugge, R., & Schoormans, J. P. L. (2012). Newer is better! The influence of a novel appearance on the perceived performance quality of products. *Journal of Engineering Design*, *23*(6), 469–484. https://doi.org/10.1080/09544828.2011.618802
110. Mugge, R., Schoormans, J. P. L., & Schifferstein, H. N. J. (2009). Emotional bonding with personalised products. *Journal of Engineering Design*, *20*(5), 467–476. https://doi.org/10.1080/09544820802698550

111. Munoz, D., & Tucker, C. S. (2016). Modeling the Semantic Structure of Textually Derived Learning Content and its Impact on Recipients' Response States. *Journal of Mechanical Design*, *138*(4), 042001-042001–042012. https://doi.org/10.1115/1.4032398
112. Nadler, G., Smith, J. M., & Frey, C. E. (1989). Research needs regarding formulation of the initial design problem. *Design Studies*, *10*(3), 151–154. https://doi.org/10.1016/0142-694X(89)90032-X
113. Neroni, M. A., Vasconcelos, L. A., & Crilly, N. (2017). Computer-Based "Mental Set" Tasks: An Alternative Approach to Studying Design Fixation. *Journal of Mechanical Design*, *139*(7), 071102-071102–071110. https://doi.org/10.1115/1.4036562
114. Newland, P., Powell, J. A., & Creed, C. (1987). Understanding architectural designers' selective information handling. *Design Studies*, *8*(1), 2–16. https://doi.org/10.1016/0142-694X(87)90026-3
115. Nikander, J. B., Liikkanen, L. A., & Laakso, M. (2014). The preference effect in design concept evaluation. *Design Studies*, *35*(5), 473–499. https://doi.org/10.1016/j.destud.2014.02.006
116. Onarheim, B., & Christensen, B. T. (2012). Distributed idea screening in stage–gate development processes. *Journal of Engineering Design*, *23*(9), 660–673. https://doi.org/10.1080/09544828.2011.649426
117. Ostergaard, K. J., III, W. R. W., Divekar, A., Vitali, H., & Summers, J. D. (2005). An experimental methodology for investigating communication in collaborative design review meetings. *CoDesign*, *1*(3), 169–185. https://doi.org/10.1080/15710880500298520
118. Perttula, M. K., Krause, C. M., & Sipilä, P. (2006). Does idea exchange promote productivity in design idea generation? *CoDesign*, *2*(3), 125–138. https://doi.org/10.1080/15710880600797942
119. Perttula, M., & Sipilä, P. (2007). The idea exposure paradigm in design idea generation. *Journal of Engineering Design*, *18*(1), 93–102. https://doi.org/10.1080/09544820600679679
120. Phillips, E. M. (1982). Understanding children's perceptions— a classroom experiment. *Design Studies*, *3*(4), 197–203. https://doi.org/10.1016/0142-694X(82)90037-0
121. Portillo, M. B., & Dohr, J. H. (1989). Design education: on the road towards thought development. *Design Studies*, *10*(2), 96–102. https://doi.org/10.1016/0142-694X(89)90042-2
122. Powell, J. A., & Newland, P. (1994). Informing multimedia: a sensitive interface to data for construction design professionals. *Design Studies*, *15*(3), 285–316. https://doi.org/10.1016/0142-694X(94)90015-9
123. Reid, F. J. M., & Reed, S. E. (2007). Conversational grounding and visual access in collaborative design. *CoDesign*, *3*(2), 111–122. https://doi.org/10.1080/15710880601143195
124. Reid, T. N., Gonzalez, R. D., & Papalambros, P. Y. (2010). Quantification of Perceived Environmental Friendliness for Vehicle Silhouette Design. *Journal of Mechanical Design*, *132*(10), 101010-101010–101012. https://doi.org/10.1115/1.4002290
125. Reid, T. N., MacDonald, E. F., & Du, P. (2013). Impact of Product Design Representation on Customer Judgment. *Journal of Mechanical Design*, *135*(9), 091008-091008–091012. https://doi.org/10.1115/1.4024724
126. Rhodes, P. J., & Powell, J. A. (1994). Tell them a story!: a theoretical model to assist in the transfer of architectural design information through video. *Design Studies*, *15*(3), 270–284. https://doi.org/10.1016/0142-694X(94)90014-0
127. Rhodes, P., McCartney, K., & Powell, J. (1992). Narrative structures within architectural videos. *Design Studies*, *13*(4), 355–378. https://doi.org/10.1016/0142-694X(92)90131-S
128. Rijn, H. van, Bahk, Y., Stappers, P. J., & Lee, K.-P. (2006). Three factors for contextmapping in East Asia: Trust, control and nunchi. *CoDesign*, *2*(3), 157–177. https://doi.org/10.1080/15710880600900561
129. Rijn, H. van, Visser, F. S., Stappers, P. J., & Özakar, A. D. (2011). Achieving empathy with users: the effects of different sources of information. *CoDesign*, *7*(2), 65–77. https://doi.org/10.1080/15710882.2011.609889
130. Robinson, M. A. (2012). How design engineers spend their time: Job content and task satisfaction. *Design Studies*, *33*(4), 391–425. https://doi.org/10.1016/j.destud.2012.03.002
131. Römer, A., Leinert, S., & Sachse, P. (2000). External Support of Problem Analysis in Design Problem Solving. *Research in Engineering Design*, *12*(3), 144–151. https://doi.org/10.1007/s001630050029
132. Sacharin, V., Gonzalez, R., & Andersen, J.-H. (2011). Object and user levels of analyses in design: the impact of emotion on implicit and explicit preference for 'green' products. *Journal of Engineering Design*, *22*(4), 217–234. https://doi.org/10.1080/09544820903158850
133. Sancar, F. H. (1996). Behavioural knowledge integration in the design studio: an experimental evaluation of three strategies. *Design Studies*, *17*(2), 131–163. https://doi.org/10.1016/0142-694X(95)00029-Q
134. Savage, J. C. D., Moore, C. J., Miles, J. C., & Miles, C. (1998). The interaction of time and cost constraints on the design process. *Design Studies*, *19*(2), 217–233. https://doi.org/10.1016/S0142-694X(98)00004-0
135. Savoie, T. B., & Frey, D. D. (2012). Detecting mistakes in engineering models: the effects of experimental design. *Research in Engineering Design*, *23*(2), 155–175. https://doi.org/10.1007/s00163-011-0120-y
136. Schaffhausen, C. R., & Kowalewski, T. M. (2015). Assessing Quality of User-Submitted Need Statements From Large-Scale Needfinding: Effects of Expertise and Group Size. *Journal of Mechanical Design*, *137*(12), 121102-121102–121109. https://doi.org/10.1115/1.4031655

137. Schütze, M., Sachse, P., & Römer, A. (2003). Support value of sketching in the design process. *Research in Engineering Design*, *14*(2), 89–97. https://doi.org/10.1007/s00163-002-0028-7
138. Seitamaa-Hakkarainen, P., & Hakkarainen, K. (2001). Composition and construction in experts' and novices' weaving design. *Design Studies*, *22*(1), 47–66. https://doi.org/10.1016/S0142-694X(99)00038-1
139. She, J., & MacDonald, E. (2013). Priming Designers to Communicate Sustainability. *Journal of Mechanical Design*, *136*(1), 011001-011001–011012. https://doi.org/10.1115/1.4025488
140. Sosa, M. E., & Marle, F. (2013). Assembling Creative Teams in New Product Development Using Creative Team Familiarity. *Journal of Mechanical Design*, *135*(8), 081009-081009–081013. https://doi.org/10.1115/1.4024763
141. Spence, C., & Zampini, M. (2007). Affective design: modulating the pleasantness and forcefulness of aerosol sprays by manipulating aerosol spraying sounds. *CoDesign*, *3*(sup1), 107–121. https://doi.org/10.1080/15710880701362679
142. Stamps III, A. E. (1999). Architectural detail, Van der Laan septaves and pixel counts. *Design Studies*, *20*(1), 83–97. https://doi.org/10.1016/S0142-694X(98)00025-8
143. Steen, M., Arendsen, J., Cremers, A., Jong, A. D., Jong, J. D., & Koning, N. D. (2013). Using interactive model simulations in co-design: An experiment in urban design. *CoDesign*, *9*(1), 2–16. https://doi.org/10.1080/15710882.2012.755549
144. Thimmaiah, S., Phelan, K., & Summers, J. D. (2017). An Experimental Study on the Influence That Failure Number, Specialization, and Controls Have on Confidence in Predicting System Failures1. *Journal of Mechanical Design*, *139*(1), 011102-011102–011112. https://doi.org/10.1115/1.4034789
145. Toh, C. A., & Miller, S. R. (2016a). Choosing creativity: the role of individual risk and ambiguity aversion on creative concept selection in engineering design. *Research in Engineering Design*, *27*(3), 195–219. https://doi.org/10.1007/s00163-015-0212-1
146. Toh, C. A., & Miller, S. R. (2016b). Creativity in design teams: the influence of personality traits and risk attitudes on creative concept selection. *Research in Engineering Design*, *27*(1), 73–89. https://doi.org/10.1007/s00163-015-0207-y
147. Toh, C. A., Miller, S. R., & Okudan Kremer, G. E. (2014). The Impact of Team-Based Product Dissection on Design Novelty. *Journal of Mechanical Design*, *136*(4), 041004-041004–041010. https://doi.org/10.1115/1.4026151
148. Toh, C. A., Strohmetz, A. A., & Miller, S. R. (2016). The Effects of Gender and Idea Goodness on Ownership Bias in Engineering Design Education. *Journal of Mechanical Design*, *138*(10), 101105-101105–101108. https://doi.org/10.1115/1.4034107
149. Toh, C., Miller, S., & Simpson, T. (2015). The impact of virtual product dissection environments on student design learning and self-efficacy. *Journal of Engineering Design*, *26*(1–3), 48–73. https://doi.org/10.1080/09544828.2015.1018149
150. Tsai, T., Chang, T.-C., Chuang, M.-C., & Wang, D.-M. (2008). Exploration in emotion and visual information uncertainty of websites in culture relations. *International Journal of Design*, *2*(2).
151. Ulusoy, Z. (1999). To design versus to understand design: the role of graphic representations and verbal expressions. *Design Studies*, *20*(2), 123–130. https://doi.org/10.1016/S0142-694X(98)00030-1
152. Valencia-Romero, A., & Lugo, J. E. (2017). An immersive virtual discrete choice experiment for elicitation of product aesthetics using Gestalt principles. *Design Science*, *3*. https://doi.org/10.1017/dsj.2017.12
153. Vasconcelos, L. A., Cardoso, C. C., Sääksjärvi, M., Chen, C.-C., & Crilly, N. (2017). Inspiration and Fixation: The Influences of Example Designs and System Properties in Idea Generation. *Journal of Mechanical Design*, *139*(3), 031101-031101–031113. https://doi.org/10.1115/1.4035540
154. Verma, N. (1997). Design theory education: how useful is previous design experience? *Design Studies*, *18*(1), 89–99. https://doi.org/10.1016/S0142-694X(96)00015-4
155. Vermillion, S. D., Malak, R. J., Smallman, R., & Linsey, J. (2015). A Study on Outcome Framing and Risk Attitude in Engineering Decisions Under Uncertainty. *Journal of Mechanical Design*, *137*(8), 084501-084501–084504. https://doi.org/10.1115/1.4030434
156. Verstijnen, I., van Leeuwen, C., Goldschmidt, G., Hamel, R., & Hennessey, J. (1998). Sketching and creative discovery. *Design Studies*, *19*(4), 519–546. https://doi.org/10.1016/S0142-694X(98)00017-9
157. Viswanathan, V., Atilola, O., Esposito, N., & Linsey, J. (2014). A study on the role of physical models in the mitigation of design fixation. *Journal of Engineering Design*, *25*(1–3), 25–43. https://doi.org/10.1080/09544828.2014.885934
158. Viswanathan, V. K., & Linsey, J. S. (2012). Physical Models and Design Thinking: A Study of Functionality, Novelty and Variety of Ideas. *Journal of Mechanical Design*, *134*(9), 091004-091004–091013. https://doi.org/10.1115/1.4007148
159. Viswanathan, V. K., & Linsey, J. S. (2013a). Design Fixation and Its Mitigation: A Study on the Role of Expertise. *Journal of Mechanical Design*, *135*(5), 051008-051008–051015. https://doi.org/10.1115/1.4024123

160. Viswanathan, V. K., & Linsey, J. S. (2013b). Role of Sunk Cost in Engineering Idea Generation: An Experimental Investigation. *Journal of Mechanical Design*, *135*(12), 121002-121002–121012. https://doi.org/10.1115/1.4025290
161. Wan, T. B. (1987). The architect as a purchaser of CAD systems: a study of potential cognitive dissonance. *Design Studies*, *8*(2), 96–101. https://doi.org/10.1016/0142-694X(87)90007-X
162. Wei, S.-T., Ou, L.-C., Luo, M. R., & Hutchings, J. B. (2014). Package design: Colour harmony and consumer expectations. *International Journal of Design*, *8*(1). Retrieved from http://search.proquest.com/openview/c687ade58e6511b44872430242ced1de/1?pq-origsite=gscholar
163. Wetmore III, W. R., Summers, J. D., & Greenstein, J. S. (2010). Experimental study of influence of group familiarity and information sharing on design review effectiveness. *Journal of Engineering Design*, *21*(1), 111–126. https://doi.org/10.1080/09544820802238217
164. Yang, M. C. (2009). Observations on concept generation and sketching in engineering design. *Research in Engineering Design*, *20*(1), 1–11. https://doi.org/10.1007/s00163-008-0055-0
165. Yang, M. C., & Cham, J. G. (2006). An Analysis of Sketching Skill and Its Role in Early Stage Engineering Design. *Journal of Mechanical Design*, *129*(5), 476–482. https://doi.org/10.1115/1.2712214
166. Yilmaz, S., Daly, S. R., Seifert, C. M., & Gonzalez, R. (2015). How do designers generate new ideas? Design heuristics across two disciplines. *Design Science*, *1*. https://doi.org/10.1017/dsj.2015.4

CHAPTER 3:
Attentional Allocation in Visual Displays for Emergencies

3.1. Chapter Introduction

In the previous chapter, deductive studies on human behaviour in design were reviewed, and the potential benefit of using theoretical and methodological insights from cognitive science was established. In alignment with this finding, this chapter presents and applies select cognitive theories of attention, and their methodological underpinnings, which could benefit design research due to the relevance of the attention system for how information is processed.

In emergencies, operators and decision makers are often required to obtain correct information at a fast pace to avoid adverse events. For example, operators of nuclear power plant control rooms need to find and report correct values pertaining to e.g. pressure or temperature as part of the diagnosing procedure. The cognitive mechanism which plays the biggest role in determining which pieces of information that enter consciousness is the attention system (Baars & Franklin, 2007; Dehaene et al., 2014; Lamme, 2003). Consequently, the design of visual interfaces would benefit from being as optimised as possible towards ensuring that users will be able to direct their attention in an efficient manner when viewing them. However, the allocation of attention is influenced by a wide variety of forces, which interplay to determine the final allocation (Bundesen & Habekost, 2008). Therefore, knowing which aspects of attention that are especially influential for any given design case is of great importance for ensuring fast and accurate performance. A large amount of research has been conducted within the various fields of attention research, which has resulted in a wide variety of topic-specific (e.g. Lavie & Tsal, 1994) and general theories of attention (e.g. Bundesen, 1990; Wolfe, 2007). As discussed previously, however, these theories are seldom considered by designers when constructing interfaces, perhaps due to the studies from which they are deduced taking place in a highly controlled laboratory environment rather than applied cases. In response to this absence of translating theory to practice, the work conducted in this PhD project on attention allocation attempted to both summarise research in attention research in a way that is more accessible to practitioners and to test predictions from attention theory in more applied contexts. Furthermore, the research was done using research methodology based on the paradigms used in experimental psychology, but adapted to better fit a more applied context so as to increase ecological validity. Through this application of attention theory and methodology, the research in this thesis produced findings that were new to attention theory, design research, and design practice.

The present chapter contains four papers that cover these aspects to varying degrees. Before outlining the original research, however, three foundational concepts are briefly outlined that are vital for understanding the attention system: *Capacity, processing* and *modularity*. To illustrate these concepts, consider a basic visual search task where an operator is attempting to find a critical piece of information in a table. In other words, what the operator is trying to accomplish is to find one specific piece of information amongst many similar pieces. The first concept, *capacity*, refers to the fact that you will only be able to inspect a few pieces of information (usually 3-4 pieces, Sperling, 1960) in your search at any given time. This constraint is central to much of attention theory, as it establishes attention's role as the allocator of sparse resources (Bundesen & Habekost, 2008). The second concept, *processing,* refers to the manner with which allocation of attention occurs. Two major types of processing bear relevancy for the presented research: The first is serial, one-by-one search (first discovered by Sternberg, 1966). In the aforementioned search task, this would equate to the operator considering one piece of information at a time, eventually stopping when the desired information is obtained. The second is search facilitated by pre-attentive processing which allows for parallel processing due to the information 'popping out' in an attribute such as colour (first discovered by Treisman & Gelade, 1980). In the visual search task, this equates to the operator scanning all pieces of information before directing attention, which allows for more efficient search. The third concept, *modularity*, refers to the fact that attentional mechanisms, including capacity and processing type, can be modulated by external factors, such as cognitive and visual load (Lavie & Tsal, 1994). This means that pieces of information that the operator is capable of processing at one time, as well as the manner with which the operator searches for information, can differ depending on the context. This is particularly important for the present thesis, as emergencies often are associated with increased cognitive load. In sum, attention is the mechanism responsible for deploying sparse resources in either a serial or parallel manner, depending on the objects that it is allocated and the context. Having established these concepts, the papers in this chapter, which all have at their basis these fundamental aspects of attention, are now outlined.

3.2. Research and Findings

The first paper (Paper 2/ Andersen & Maier, 2016) outlines the most important object characteristics that capture our gaze and factors that modulate this capture. Furthermore, the paper offers suggestions for how these findings could be used by designers to create visual interfaces that more efficiently deploy attention. The characteristics included were: Emotion, contrast, meaningfulness, faces, onset of new stimuli, and task-relevance. The modulating factors included were mental load (previously referred to in this thesis as cognitive load), visual load, and alertness. By outlining these factors, the paper serves as a basis for both practitioners and design researchers who seek to understand where people direct their gaze, due to the relation between attentional allocation and gaze direction.

The second and third paper (Paper 3/ Andersen & Maier, 2017; Paper 4/ Andersen & Maier, submitted) investigated the extent to which six colours (red, blue, yellow, green, purple, orange) can guide attention in a complex display. Previous research has shown that colour can guide and capture attention (Wolfe & Horowitz, 2017), but thus far no studies have investigated whether specific colours guide attention differently. Therefore, we created a visual search paradigm where participants search for a specific coloured circle amongst 19 other coloured circles, indicating whether the target circle is present or not. A display could contain one coloured target circle and between one and four other coloured circles. Both studies showed different guidance and capture for specific colours, with red being found consistently fastest and orange and purple being found consistently slower. Andersen & Maier (submitted) expanded upon the sample size and analysis of Andersen & Maier (2017), measuring whether the difference in individual guidance interacted with the visual complexity, measured as the amount of different colours present in the display. Consistent with predictions from Load Theory (Biggs et al., 2015), it was found that the difference in guidance increased as complexity increased. The results have implication for both attention theory, which thus far has only considered colour as a category, rather than studying specific colours, and for application, where the use of colour is common practice as a means for improving performance.

The fourth paper (Paper 5/ Andersen et al., submitted) investigated which of a series of compatible and contradictory predictions from attention theory would hold in a complex display with objects that resemble the objects likely to be found in a real life nuclear control room display. As with the two previous papers, the study investigated design relevant predictions from attention theory in a setting that was more applied than usual attention research paradigms, but was controlled enough to allow testing of specific theoretical predictions. Specifically, the study tested predictions from Feature Integration Theory (Treisman & Gelade, 1980), Load Theory (Lavie & Tsal, 1994) and Dilution Theory (Benoni & Tsal, 2010; Tsal & Benoni, 2010), as well as specific predictions about the attentional guidance of colour (Paper 3/ Andersen & Maier, 2017; Paper 4/ submitted; Biggs et al., 2015; Wolfe & Horowitz, 2017). A new visual search task was created based on the objects found in a nuclear power plant control room aligned in a manner that mimicked the experimental paradigms of the aforementioned attention studies. This allowed for testing specific predictions of the theories in a more complex and realistic display. The results showed that several predictions from attention theory could generalise to a more complex display. Notably, the prediction of Feature Integration Theory regarding search under high and visual load was accurate for both experiments, which was reflected in the fact that irrelevant distracting objects did not slow affect search. These findings, as with the previous studies thus have implication for designing of visual interfaces that are aligned with the natural tendencies of our attentional system. Furthermore, our results give cadence to a recent critique of Load Theory (Benoni & Tsal, 2010; Kyllingsbæk et al., 2011), and thus have a concrete impact on attention theory as well.

3.3. Implications for Design Practice

The four papers presented in this chapter have implications for how visual interfaces are designed, as guidelines are given based on attention theory, and tests of attention theory, on how attention is expected to be allocated. Designers looking to create complex visual interfaces for emergencies can apply the outlined theory and research findings with regards to the role of colour, cognitive load, visual load, visual dilution and the other factors reviewed in Andersen & Maier (2016). Through these insights, designer could for example increase the likelihood that critical information is seen by operators using complex interfaces by using colours that match the importance of the objects with their attentional guidance and capture. Based on the findings in Paper 3 and 4 (Andersen & Maier, 2017; submitted), the highest priority object could thus be coloured red, whereas the least important object could be coloured purple. Furthermore, if several objects have equal importance, our findings could likewise guide designers to choosing colours that affect attention equally. The findings in Paper 4 and 5 (Andersen & Maier, submitted; Andersen et al., submitted) show that that this alignment is particularly important for the design of complex visual interfaces, as the observed differences in guidance by colour changed based on the visual load as induced by differences in the complexity of the display.

Paper 2:

What Captures Gaze in Visual Design? Insights from Cognitive Psychology (2016), by Andersen, E., & Maier, A. Proceedings of NordDesign 2016, Volume 1, Trondheim, Norway, 10th-12th August 2016

What Captures Gaze in Visual Design? Insights from Cognitive Psychology

Emil Andersen[1], Anja Maier[1]

[1]Technical University of Denmark
emand@dtu.dk
[2] Technical University of Denmark
amai@dtu.dk

Abstract

Visual information is vital for user behaviour and thus of utmost importance to design. Consequently, tracking and interpreting gaze data has been the target of increasing amounts of research in design science. This research is in part facilitated by new methods, such as eye-tracking, becoming more readily available. Visual attention is the principle mechanism that governs where we direct our gaze. Understanding the factors that influence how attention is directed is therefore necessary for understanding user intentions and gaze patterns.

In this paper, we provide an overview of the characteristics and factors that have been experimentally shown to capture attention, as well as those factors that modulate the capture and direction of attention. We do so by drawing on the large body of evidence provided by cognitive psychology, as we believe this research area could potentially provide a source of untapped potential for design research and practice.

Keywords: Attention, Design, Gaze, Eye-tracking, Visual Communication

1. Introduction

Visual information is of utmost importance when designing products or interfaces for both leisure and labour, as users judge the perceived value based on appearance of the products or interfaces (Crilly et al., 2004). Consequently, visual processing has increasingly been the target of research within design science, especially as novel methods for tracking the visual operations of users have become more readily available. In particular, the use of eye-tracking devices has shown increasing promise, as such devices allow for real-time insights into the direction of the gaze. Examples include a wide range of topics such as: Navigating a website (Goldberg et al., 2002; McCarthy et al., 2004), reading a diagram for use in work (Störrle et al., 2014; Lohmeyer et al., 2015; Maier et al., 2015) or monitoring how operators perform in realistic simulations of high-stakes situations in a chemical laboratory (Sharma et al., 2016). Through such investigations of the user's gaze, researchers and designers alike have found a powerful tool for shaping products and interfaces in ways that allow users to operate in accordance with both their own and the producer's intentions.

However, interpreting the reasons behind gaze movements is not always a straightforward task. Research, in particular from the field of cognitive psychology, has shown that the direction of our attention, and consequently our gaze, is not always under our control (Theeuwes et al., 1998). In this paper, we highlight characteristics that have been empirically investigated though rigorous cognitive experiments, and have been confirmed to *capture* attention. They are, in no particular order: Emotion (e.g. Öhman et al, 2001a; 2001b), contrast (usually colour, e.g. Treisman & Gelade, 1980; Nordfang, 2013), meaningfulness (e.g. Biggs et al., 2012), faces (e.g. faces Ro et al., 2001), onset of new stimuli (Jonidese, 1981; Theeuwes, 1990; Theeuwes et al., 1998) and task-relevance (Hodsoll et al., 2011). Furthermore, we discuss external factors that impact how well we are able to direct our attention in accordance with our goals and to ignore distractors (Lavie & Tsal, 1994; Lavie et al., 2004; Geng, 2014). Using psychological principles to form the basis of design science has led to many fruitful insights in the past: For example, the gestalt laws and SRK (skills, rules, knowledge) principle, which is based on psychological insights, has informed design decisions of displays for the nuclear and oil industries (Lau et al., 2008; Weyer et al., 2010; Braseth & Øritsland, 2013; Lau et al., 2008). Another example concept of functional fixedness (Duncker & Lees, 1945) has inspired a diverse and informative range of studies of what is now known as design fixation (Jansson & Smith, 1991; Purcell & Gero, 1996; Vasconecelos & Crilly, 2016). The present paper similarly aspires to inspire design research and practice: The papers selected in this overview seminal works in cognitive psychology have in common that they a) are extremely rigorous, and b) work under the assumption that, given the extreme rigour, their findings reflect basic psychological functions that reflect the way our brain is wired (and thus should be similar in all humans). In response to calls and movements for higher degrees of rigour in experimental design science (Papalambros, 2015; Cash et al., 2016), we hope that presenting these studies will inspire similar studies, as well as provide insights into how to interpret how gaze is captured.

These aims are accomplished in four parts. First, we establish some basic fundamental aspects of attention for use in further discussion. Second, we describe visual features that have been experimentally shown to capture attention. Third, we establish factors that mediate the extent of this attentional capture. Fourth and finally, we discuss potential implications of the presented insights for designers and design researchers alike.

2. Attention: Conceptual Fundamentals from Psychology

Given the complexity and richness of the stimuli in our surroundings – be it the pages of a website or the dashboards of a control room – a key prerequisite for interacting with products and interfaces is our ability to discriminate between relevant and irrelevant items, and to focus on the former while ignoring the latter. The processes and mechanisms that allow these selective operations are collectively referred to as attention, or selective attention (e.g. Deutsch & Deutsch, 1963). Furthermore, attention serves as a principle precursor for the movement of our gaze (Hoffman & Subramaniam, 1995; Kowler et al., 1995; Theeuwes et al., 1998), and we therefore largely frame our discussion around attention research.

Despite our frequent use and sense of intuitive understanding of the concept and despite decades of research on the topic (Posner & Petersen, 1990; Bundesen & Habekost, 2008), attention has so far remained elusive in experimental psychology and it is a highly debated and researched topic in all variations of cognitive psychology – even to this date. However, two fundamental properties of attention have been established that are important for this paper: One relating to the limited capacity of attention, and another relating to how stimuli are selected when several are present. Each is considered in the following as they provide a basis for understanding the subsequent sections of this paper.

2.1. Attention has Limited Capacity

Given that attention is primarily intended as a function that selects some information while ignoring others, it follows naturally that the capacity of the attentional system be substantially limited in comparison to the vast number of stimuli in our surroundings. In humans, the process of focusing is generally extremely successful, and we are consequently only able to hold 3-4 objects in our visual attention at any given time (Sperling, 1960). However, this capacity may be increased by "chunking" objects together to form objects with larger amounts of information (Miller, 1956) such as when remembering an 8 single-digit number as 4 two-digit numbers. Nevertheless, the principal fundamental feature of the attention system is that it has limited capacity.

2.2. Attention is Directed and Captured by Competing Stimuli

The prevailing view on the way stimuli are selected for one of the limited slots of attentional capacity is that of Biased Competition (Desimone & Duncan, 1995). In this theoretical framework, when viewing a scene, our attention system assigns relevance scores to stimuli depending on their characteristics. Following this, stimuli with higher relevance scores are more likely (but not guaranteed) to win the competition and be selected by attention and thereby allowed further processing. The process has sometimes been envisioned as a stochastic race (for details on this model, see Bundesen (1990) and Bundesen & Habekost (2008)). The Biased Competition framework furthermore serves as a powerful metaphor for understanding gaze direction; gaze-direction should be seen as inherently random, but skewed towards certain characteristics depending on the circumstances. In this light, the purpose of this paper is to summarise those characteristics that have been experimentally proven to consistently bias the competition towards certain gaze directions.

3. Characteristics that Capture Attention

Having defined some fundamentals of how attention works, we now turn to those characteristics that have been shown experimentally to capture attention. Possible confounding factors are listed and discussed where relevant. The characteristics are: Contrast, emotion and meaning, faces, onset of new stimuli and task-relevance. Outlining these characteristics serve as insights to consider when interpreting gaze direction. Factors that mediate the characteristics' attentional capture are discussed in section four of the paper.

3.1. Contrast

As possibly the earliest characteristic that was identified to capture attention, the effect of contrast is observed when one object deviates substantially from its surroundings in colour or shape. While other physical features, such as gestalt laws (Moore & Fitz, 1993), aesthetics (Crilly et al., 2004) or concrete layout features (e.g. number of line crossings in diagrams, Maier et al., 2014) have proven relevant for design decisions, contrast, in particular between colours, has been especially important for, and well investigated by, experimental psychologists. For example: When searching for a target letter amongst a display of letters, a coloured singleton presented alongside letters of a different colour will become immediately available for attention. This finding was a fundamental part of one of the earliest theories of attention (The Feature Integration Theory, Treisman & Gelade, 1980). This finding has been reproduced in numerous designs, and, importantly, has been reproduced in a context where colour is not a target-relevant criterion (Forster & Lavie, 2008b; Nordfang et al., 2013, see the importance of separating the feature effect from task-relevance below). Objects that differ from their surroundings in colour contrast thus represent one of the most robust and important examples of attentional capture in experimental psychology.

3.2. Emotion and Meaning

Stimuli that carry emotional value –be it positive (e.g. happiness) or negative (e.g. anger, threat) – have been extensively studied and shown to capture attention. Early studies focused on phylogenetic threat, showing that a threatening singleton (spider, snake) is found more rapidly amongst neutral distractors (mushrooms, flowers) than vice versa (Öhman, 2001a). Other studies have expanded on this research by showing that ontogenetic threats (i.e. a gun or knife: Brosch & Sharma, 2005; Blanchette, 2006), angry faces (Hansen & Hansen, 1988; Öhman et al., 2001b; Calvo et al, 2006), but also positive stimuli such as faces (Nummenmaa et al., 2006) or pictures of emotional scenes (Calvo & Lang, 2004) capture attention.

However, recent studies have questioned whether these effects are due to emotion specifically, or whether the effects were caused by task-relevance (see below) or simply that emotion was often the only meaningful category. A study by Biggs et al., (2012) found that meaningfulness in itself is indeed enough to capture attention, but some evidence has suggested that emotion may capture attention because of their valence and not just their meaning (Öhman, 2001a). However, this debate is largely unresolved and researchers are therefore advised to carefully consider the relative importance of emotion and meaning in their studies.

3.3. Faces

Identifying the facial expression of our fellow humans is vital to everyday life. This is reflected in a dedicated area for facial processing in the brain (the Fusiform Face Area; Kanwisher, 1997), but also in attentional capture. We previously highlighted studies that show that emotional faces capture attention. However, faces in general are processed more efficiently, are more easily detected and capture attention (Ro et al., 2001; Theeuwes & Van der Stigchel, 2006; Langton et al., 2008).

3.4. Onset of New Stimuli

The onset of new stimuli in the visual field has continuously been shown to capture attention and to direct the gaze (Theeuwes, 1991; Hillstrom & Yantis, 1994; Theeuwes et al., 1998; Cosman & Vescera, 2009). In fact, the appearance of new stimuli has such powerful attentional capture that some researchers have proposed that they are categorically "special" in how they are treated by the attention system (Theeuwes, 2010; Nordfang & Bundesen, 2010). Using this mechanism for attentional capture should consequently be done with care.

3.5. Task Relevance

Task relevance is an important factor in determining how likely an object is to capture attention. For instance, Nordfang et al., (2013) showed that objects with high task relevance were more likely to capture attention, and that task-relevance interacted additively with contrast such that targets of high contrast and relevance were even more likely to capture attention than stimuli that held only one of the characteristics. Similar findings have been shown for emotional stimuli (Hodsoll et al., 2011). Importantly, these studies have in common that they inform the participant about the target characteristics, which in turn leads to facilitated search for these task relevant objects. It may therefore be that objects that conform to the usual appearance of an object with a specific function are more readily detected due to their task-relevance (as opposed to any of the above factors) (Geng, 2014).

A possible complication with attentional capture due to task relevance lies in whether only low complexity stimuli such as letters and numbers may lead to this capture (Nordfang et al., 2013). However, recent evidence has shown that also high-complexity stimuli such as pictures can guide attention if they are task relevant (Alexander & Zelinsky, 2012).

3.6. Interactions between Characteristics

The relative importance or strength of the highlighted characteristics has so far not been well investigated. However, evidence exists for some examples of interactions: First, if several faces are present in the visual field, emotional faces will capture attention over neutral faces, and faces with negative emotion will capture attention over faces with positive emotion (Hansen & Hansen; Öhman et al., 2001b; Nummenmaa et al., 2006). Second, as noted above, if targets are task relevant and differ from its surroundings in colour contrast, the search for this target will be faster if the target has only one of the characteristics (Nordfang et al., 2013). This synergy is likely to be present between task relevance and other attention-capturing characteristics, but to our knowledge this has not been investigated experimentally. Similarly, it is possible that other synergistic relationships exist, as is the case with emotional faces, but, again, experimental evidence has thus far been lacking. As interfaces and products often comprise multiple and complex stimuli, determining these relationships could be vital for understanding the direction of gaze, and consequently provides an interesting target for research.

4. Factors that Facilitate and Inhibit Attentional Control

Having outlined characteristics that are highly likely to capture attention, we now focus on external factors that influence our ability to direct attention and ignore irrelevant stimuli. Importantly, the presented factors should be considered as interacting with the characteristics outlined in section 3 and eventually deciding direction of the gaze.

4.1. Cognitive Load

In their influential theory of attention, Lavie and Tsal (1994, see also Lavie et al., 2004) describe the relationship between likelihood that distractors affect behaviour and cognitive load. They propose two mechanisms: The first mechanism relates to the capacity constraint of attention and states that if the visual field contains few target objects, then the leftover capacity will be filled by distractors until the capacity is reached. This is in line with the outlined Biased Competition framework: If there is no competition, then everyone is a winner.

The second mechanism is that high mental load, such as from other tasks or the environment, will reduce our ability to suppress distractors because less mental capacity will be available for this suppression. The theory thereby establishes that high workload is associated with increased distractor interference due to lacking inhibition, while low workload may lead to distractor processing if the number of relevant objects is below the total capacity.

4.2. Alertness and Mental Fatigue

An important function of attention, which was most notably reviewed by Posner & Petersen (1990; Petersen & Posner, 2012), lies in preparing and sustaining alertness for high priority stimuli. This alertness, they noted, is reflected in faster response times, but sometimes comes at the cost of higher error rates because responses are made based on less information. For instance, relating to gaze direction, this could mean that users who are focused on finding a certain feature in an interface may direct their gaze rapidly to features that are congruent with their intentions, but that they may commit erroneous gaze directions if many similar objects are present in the visual field.

More recently, Boksem et al., (2005) found that mental fatigue results in symptoms that are similar to this state: As participants worked on a tiring visual attention task for 3 hours, performance decreased steadily due to increased reaction times and distractibility, as well as more false alarms and errors. This is corroborated by Geng (2014) who found that the proactively suppressing distracting information is mentally demanding - and next to impossible if very little is known about the target's characteristics. Therefore, she notes, we usually rely on reactive corrections of initial misfires of attention. In other words, we let ourselves get distracted and correct subsequently, because proactively suppressing distractors is mentally strenuous, which further underlines the importance of knowing where the gaze goes initially.

Table 1. Overview of Characteristics and Factors that Influence the Capture of Gaze

		Description	References from Cognitive Psychology
Characteristics	Contrast	Differences in physical features, such as colour or shape	Treisman & Gelade, 1980; Forster & Lavie, 2008b; Nordfang et al., 2013
	Emotion	Differences in emotional valence	Öhman et al., 2001a; 2001b; Brosch & Sharma, 2005; Blanchette, 2006: Hansen & Hansen, 1988; Calvo et al., 2006; Nummenmaa et al., 2006; Calvo & Lang, 2004
	Meaning	Being meaningful rather than meaningless	Biggs et al., 2012
	Faces	Pictures of human faces or stylistic drawings	Kanwisher, 1997; Ro et al., 2001; Theeuwes & Van der Stigchel, 2006; Langton et al., 2008
	Onset	Onset of new stimuli	Theeuwes, 1991; Hillstrom & Yantis, 1994; Theeuwes et al., 1998; Cosman & Vescera, 2009; Theeuwes, 2010
	Task-Relevance	Degree of congruence with the target of a task	Hodsoll et al., 2011; Nordfang et al., 2013; Geng, 2014
Factors	Mental Load	High mental load interferes with inhibitory mechanisms	Lavie & Tsal, 1994; Lavie et al., 2004
	Visual Load	More relevant targets allows for less processing of distractors	Lavie & Tsal, 1994; Lavie et al., 2004
	Alertness	Alertness allows for higher sensitivity to targets, but at the cost of more errors	Posner & Petersen, 1990; Petersen & Posner, 2012; Boksem et al., 2005; Gent, 2014

5. Implications and Guidelines for Design

This paper described characteristics and factors that influence the direction of attention and thereby the gaze. We presented evidence that contrast, emotion, meaningfulness, novelty, faces and task-relevance are characteristics that capture attention. Furthermore, we offered a number of theoretical and empirical insights that shows how mental- and visual load, alertness and fatigue are factors that influence our ability to control our attention and ignore distractors thus mediating the attentional capture of the characteristics.

Based on these findings, we propose that researchers and designers should interpret the direction of the gaze to products and interfaces with care whenever these factors are involved, as they may capture attention regardless of their task relevance. Concluding this paper, we further expand how these insights may impact how researchers and designers should interpret gaze movements.

5.1. Implications for Research

For researchers, we offer that the highlighted characteristics and factors should be considered when studying the gaze in order to avoid misleading interpretations of participants' behaviour. This is especially important given that the highlighted characteristics and factors are present in almost all studies of gaze direction. Consequently, knowing whether a participant directed his/her gaze intentionally or due to attentional capture is of critical importance for interpreting results of gaze movements. In particular, task-relevance and meaningfulness were found to be important confounding factors for studies of attentional capture. Researchers should therefore carefully consider the role of these variables in their study designs.

Furthermore, we hope that future research may further improve our understanding of the exact mechanisms of user perception in design, as well as shaping the future of design. In particular, as noted above, additional research on the interplay between attention-capturing characteristics and factors is needed for truly understanding the direction of gaze.

5.2. Implications for Designers

For designers, we hope that the provided overview of characteristics and factors, as well as the research behind, can be translated directly into design practice. For instance, designers could use attentional capturing characteristics for those features that should be viewed immediately. Furthermore, designers should consider whether the conditions in which their products will be viewed will impose constraints on mental workload or fatigue, as goal directed behaviour is affected When, for example, mental load is higher, characteristics that capture attention may do so more effectively, and gaze will consequently be directed to objects with such characteristics more readily.

We envision several situations where these insights could have direct practical implications for design. For instance, planned future studies our research group will investigate how design solutions can alleviate the complexity of critical situations in control rooms, such as in those operated in a nuclear reactor or the cockpit of an airplane. One avenue of this research will be on how information processing of complex interfaces could be aided by applying the attention capturing mechanisms for highlighting the most important areas. Similar application areas should be found in all aspects of design where visual information processing is required. We therefore hope that researchers and designers alike will join us in mapping the mechanisms and design applications of insights in the visual attention system.

References

Alexander, R. G., & Zelinsky, G. J. (2012). Effects of part-based similarity on visual search: The Frankenbear experiment. Vision Research, 54, 20–30.

Biggs, A. T., Kreager, R. D., Gibson, B. S., Villano, M., & Crowell, C. R. (2012). Semantic and affective salience: The role of meaning and preference in attentional capture and disengagement. Journal of Experimental Psychology: Human Perception and Performance, 38(2), 531–541.

Braseth, A. O., & Øritsland, T. A. (2013). Visualizing complex processes on large screen displays: Design principles based on the Information Rich Design concept. Displays, 34(3), 215–222.

Brosch, T., & Sharma, D. (2005). The Role of Fear-Relevant Stimuli in Visual Search: A Comparison of Phylogenetic and Ontogenetic Stimuli. Emotion, 5(3), 360–364.

Bundesen, C. (1990). A theory of visual attention. Psychological review, 97(4), 523.
Bundesen, C., & Habekost, T. (2008). Principles of Visual Attention:: Linking Mind and Brain. Oxford University Press.
Calvo, M. G., & Lang, P. J. (2004). Gaze Patterns When Looking at Emotional Pictures: Motivationally Biased Attention. Motivation and Emotion, 28(3), 221–243.
Calvo, M. G., Avero, P., & Lundqvist, D. (2006). Facilitated detection of angry faces: Initial orienting and processing efficiency. Cognition & Emotion, 20(6), 785–811.
Cash, P., Stanković, T., & Štorga, M. (Red.). (2016). Experimental Design Research. Cham: Springer International Publishing.
Crilly, N., Moultrie, J., & Clarkson, P. J. (2004). Seeing things: consumer response to the visual domain in product design. Design Studies, 25(6), 547–577.
Desimone, R., & Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. Annual Review of Neuroscience, 18(1), 193–222.
Deutsch, J. A., & Deutsch, D. (1963). Attention: Some Theoretical Considerations. Psychological Review, 70(1), 80–90.
Duncker, K., & Lees, L. S. (1945). On problem-solving. Psychological monographs, 58(5).
Forster, S., & Lavie, N. (2008a). Attentional capture by entirely irrelevant distractors. Visual Cognition, 16(2/3), 200–214.
Forster, S., & Lavie, N. (2008b). Failures to ignore entirely irrelevant distractors: The role of load. Journal of Experimental Psychology: Applied, 14(1), 73–83.
Geng, J. J. (2014). Attentional mechanisms of distractor suppression. Current Directions in Psychological Science, 23(2), 147–153.
Goldberg, J. H., Stimson, M. J., Lewenstein, M., Scott, N., & Wichansky, A. M. (2002). Eye Tracking in Web Search Tasks: Design Implications. I Proceedings of the 2002 Symposium on Eye Tracking Research & Applications (s. 51–58). New York, NY, USA: ACM.
Hansen, C. H., & Hansen, R. D. (1988). Finding the face in the crowd: an anger superiority effect. Journal of personality and social psychology, 54(6), 917.
Hodsoll, S., Viding, E., & Lavie, N. (2011). Attentional capture by irrelevant emotional distractor faces. Emotion, 11(2), 346–353.
Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. Perception & Psychophysics, 57(6), 787–795.
Jansson, D. G., & Smith, S. M. (1991). Design fixation. Design studies, 12(1), 3–11.
Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. The Journal of Neuroscience, 17(11), 4302–4311.
Lau, N., Veland, Ø., Kwok, J., Jamieson, G. A., Burns, C. M., Braseth, A. O., & Welch, R. (2008). Ecological Interface Design in the Nuclear Domain: An Application to the Secondary Subsystems of a Boiling Water Reactor Plant Simulator. IEEE Transactions on Nuclear Science,
Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load Theory of Selective Attention and Cognitive Control. Journal of Experimental Psychology: General, 133(3), 339–354.
Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. Perception & Psychophysics, 56(2), 183–197.
Lohmeyer, Q., Mussgnug, M., & Meboldt, M. (2015). Skimming and Scrutinizing: Quantifying Two Basic Patterns of Visual Behavior in Design. I A. Chakrabarti (Red.), ICoRD'15 – Research into Design Across Boundaries Volume 1 (Bd. 34, s. 479–489). New Delhi: Springer India.
Maier, A., Baltsen, N., Christoffersen, H., & Störrle, H. (2014). Towards Diagram Understanding: A Pilot Study Measuring Cognitive Workload Through Eye-Tracking. Proceedings of International Conference on Human Behaviour in Design 2014.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychological Review, 63(2), 81–97.
Moore, P., & Fitz, C. (1993). Gestalt Theory and Instructional Design. Journal of Technical Writing and Communication, 23(2), 137–157.
Nordfang, M., & Bundesen, C. (2010). Is initial visual selection completely stimulus-driven? Acta Psychologica, 135(2), 106–108.
Nordfang, M., Dyrholm, M., & Bundesen, C. (2013). Identifying bottom-up and top-down components of attentional weight by experimental analysis and computational modeling. Journal of Experimental Psychology: General, 142(2), 510–535.
Nummenmaa, L., Hyönä, J., & Calvo, M. G. (2006). Eye movement assessment of selective attentional capture by emotional pictures. Emotion, 6(2), 257–268.
Ohman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. Journal Of Experimental Psychology-General, 130(3), 466–478.
Ohman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited:A threat advantage with schematic stimuli. Journal Of Personality And Social Psychology, 80(3), 381–396.
Papalambros, P. Y. (2015). Design Science: Why, What and How. Design Science, 1.
Posner, M. I., & Petersen, S. E. (1990). The Attention System of the Human Brain. Annual Review of Neuroscience, 13(1), 25–42.
Purcell, A. T., & Gero, J. S. (1996). Design and other types of fixation. Design Studies, 17(4), 363–383.
Ro, T., Russell, C., & Lavie, N. (2001). Changing Faces: A Detection Advantage in the Flicker Paradigm. Psychological Science, 12(1), 94–99.
Sharma, C., Bhavsar, P., Srinivasan, B., & Srinivasan, R. (2016). Eye gaze movement studies of control room operators: A novel approach to improve process safety. Computers & Chemical Engineering, 85, 43–57.
Sperling, G. (1960). The Information Available in Brief Visual Presentations. Psychological Monographs: General and Applied 1960, 74(11), 1–29.
Störrle, H., Baltsen, N., Christoffersen, H., & Maier, A. (2014). On the Impact of Diagram Layout: How Are Models Actually Read? Poster Session and ACM SRC of MODELS 2014, 31–35.
Theeuwes, J. (1991). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. Perception & psychophysics, 49(1), 83–90.
Theeuwes, J. (2010). Top–down and bottom–up control of visual selection. Acta Psychologica, 135(2), 77–99.
Theeuwes, J., Kramer, A. F., Hahn, S., & Irwin, D. E. (1998). Our Eyes do Not Always Go Where we Want Them to Go: Capture of the Eyes by New Objects. Psychological Science, 9(5), 379–385.
Theeuwes, J., & Van der Stigchel, S. (2006). Faces capture attention: Evidence from inhibition of return. Visual Cognition, 13(6), 657–665.
Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. Cognitive Psychology, 12(1), 97–136.
Vasconcelos, L. A., & Crilly, N. (2016). Inspiration and fixation: Questions, methods, findings, and challenges. Design Studies, 42, 1–32.
Weyer, U., Braseth, A. O., Eik\a as, M., Hurlen, L., Kristiansen, P., Kvalem, J., & others. (2010). Safety presentation in large screen displays-a new approach. I SPE Intelligent Energy Conference and Exhibition. Society of Petroleum Engineers.

Paper 3:

The Attentional capture of colour in visual interface design: a controlled-environment study (2017), by Andersen, E., & Maier, A. Proceedings of the 21st International Conference on Engineering Design (ICED 17) Vol 8: Human Behaviour in Design, Vancouver, Canada, 21-25.08.

THE ATTENTIONAL CAPTURE OF COLOUR IN VISUAL INTERFACE DESIGN: A CONTROLLED-ENVIRONMENT STUDY

Andersen, Emil; Maier, Anja
Technical University of Denmark, Denmark

Abstract

The use of colour is an integral component in visual interface design for creating separation between objects and for conveying meaning. It has previously been established that colours can be separated in a hierarchy of primary colours and secondary colours, and that colours are consistently associated with specific mood tones. However, it has thus far not been investigated whether these two factors, which we refer to as the perception-primacy and emotion-conveyance, are associated with attentional capture in a congruent manner. To investigate this, we conducted a visual search task study in a controlled environment, in which 11 participants scanned a 20 item display for a coloured target amongst coloured distractors. We found evidence to support that primary colours capture attention significantly more than secondary colours, and inconclusive evidence that colours convey their meaning at a sufficiently early level of processing to influence attention. We end by discussing implications of our results for design practice and research in psychology.

Keywords: Attention in design, Emotional design, Communication, Human behaviour in design, Visualisation

Contact:
Emil Andersen
Technical University of Denmark
Department of Management Engineering
Denmark
emand@dtu.dk

Please cite this paper as:
Surnames, Initials: *Title of paper*. In: Proceedings of the 21st International Conference on Engineering Design (ICED17), Vol. 8: Human Behaviour in Design, Vancouver, Canada, 21.-25.08.2017.

1 INTRODUCTION

The design of visual interfaces has become increasingly important as the prevalence of computer technology widens. For example, sectors with control rooms, such as the nuclear, are moving away from physical levers, barometers and switches towards computer-integrated systems with computer screens as the main access and information point (Lau et al., 2008; Braseth and Øritsland, 2013). In such contexts, or in everyday life, designs that allows for swift and effective behaviour are essential for performance. To achieve this, designers must create interfaces where the visual objects are clearly distinguished and easy to find. Research has shown that our ability to see in colours is particularly efficient for distinguishing between objects (Vazques et al., 2010). From an evolutionary perspective, the ability to see colour has allowed us many advantages, beyond the aesthetic, such as easily finding wild berries in a bush (see Figure 1). In the same vein, designers may use colour to make objects easily detectable in a crowded interface.

In this paper, we present a study of the interactions between colours and attention to see which colours best facilitate search tasks, and to aid designers in a world of visual interfaces. We measured attentional capture as a function of reaction time in a visual search task with regards to what we call the perception-primacy account, which states that there is a set of primary and secondary colours that cause differences in attentional capture, and the emotion-conveyance account, which states that colours convey emotions at an early enough stage to influence attentional capture.

This paper consists of five sections: Section one introduced the topic. Section two discusses the perception-primacy and emotional-conveyance accounts. Section three describes the methods of the controlled-environment experiment that we conducted to test the accounts. Section four shows the statistical analysis of the results. Section five discusses the results and suggests impacts for design and psychology.

2 BACKGROUND

2.1 Colour in the Eye and in Art

The colours we see are the result of white light being reflected on surfaces of the objects we are looking at. Differences in colour arise from differences in the wavelengths of lights that the various objects absorb: a blue object absorbs all other wavelengths than blue, a yellow object absorbs all wavelengths other than yellow etc.. White is the colour shown when light is perfectly reflected, whereas black is the colour that arises when all wavelengths are absorbed. Our retinas contain a myriad of photoreceptive cells commonly referred to as rods and cones. Of these, cones are selectively responsive to either red, green, or blue wavelengths of light. The reactions of these cells are combined by the brain to what we experience as colour (Gazzaniga, 2009; Ware, 2010).

Figure 1. Finding berries with and without colour vision

In art and design, colours are typically divided into primary and secondary colours and arranged in what is commonly referred to as a colour wheel such as the one shown in Figure 2. The primary colours (red, blue and yellow) are referred to as such due to the properties of pigments: When combining pigments, one will progressively move towards a colour that absorbs more wavelengths, ending in black. However, one can produce all the colours of the rainbow by combining red, blue and yellow in the correct manner. Conversely, none of these three colours can be achieved by combining any other sets of colours, thereby establishing them as primary. Secondary colours are then those colours that are immediate derivatives of mixing the primary colours (Harkness, 2005). While the pigment-based definition is more prevalent, it is to be underlined that other definitions of what is primary and secondary colours are equally viable, as many systems (including, as mentioned above, our visual system) is able to create all colours through a combination of red, blue and green light.

2.2 Colour and Emotion

Beyond the direct use for separating objects, colours provide a powerful way for visual designers to convey emotion and meaning. Early studies from psychology have verified that colours can reliably be associated with certain mood tones: Red is most frequently associated with exciting, defending and defying/hostile moods, blue with soothing and secure, yellow and orange with stimulating, cheerful and exciting moods, green with calm and soothing and purple with dignified and stately moods (Wexner, 1954; Murray and Deabler, 1957; Schaie, 1961a, 1961b). More recently, studies from the design literature have shown that green, blue and white colours are more effective at signalling environmental friendliness for cars than were red and black (Lee et al., 2015) and that white is perceived as being more elegant than other colours (Na and Suk, 2014). Another avenue for conveying meaning lies in typical use of colours: For example, red is used for "stop" in traffic lights and is used in warning signs, whereas green means "go" in traffic and is used for recycling badges. The efficacy of using colour for emotional conveyance is further supported by in-company research (Gillet, 2014). Applying colours that fit with the desired conveyed meaning or that induce the desired mood is thus crucial and common among designers to use when creating a visual profile (Page et al., 2012). Furthermore, research from the design community has established the importance and prevalence of colour in creating designs for mood states (Desmet, 2015).

However, the level of processing required for the conveying of this meaning has thus far yet to be investigated. Is the emotion conveyed at an early stage, before attention is directed, or only after conscious processing has been applied? This question is particularly interesting, as other emotional stimuli have been shown to capture attention at an early stage of processing and thereby to capture attention. For example, Öhman et al. (2001) found that snake and spider stimuli captures attention more than flower and mushroom stimuli, which was later echoed by Brosch and Sharma (2005), who found equivalent results with modern stimuli (guns and syringes vs cups and mobile phones). Similar effects have been produced in numerous studies, albeit with varying effect sizes depending on the specific stimuli used (e.g. Koster et al., 2004; Huang et al., 2008; Hodsoll et al., 2011). Should colours convey their meaning at an early stage of processing as well, they would therefore be expected to direct attention in a manner congruent to their associated emotion.

2.3 The Attentional Capture of Colour

Thus far, studies have focused only on whether colours capture attention at an early stage, showing different results depending on the experimental setup (Folk et al., 1994; Theeuwes, 1994; Müller et al., 2009). However, based on the evidence described above, we propose two separate accounts that, if they hold true, would result in differences in performance in a visual search task: The perception-primacy account and the emotion-conveyance account.

The perception-primacy account states that there is, as argued above, a set of primary and secondary colours, and of that colours in a higher hierarchy will be treated preferentially in the attention-system. We investigate here both whether the biologically and artistically founded notions of primary colours. Should the perception-primacy account hold true, red, green and yellow or red, green and blue (depending on whether one adopts the biological primary colours or artistic primary colours respectively) should be treated preferentially and thus capture attention more. In turn, this would be visible through faster search times for these colours when they are the target of a visual search task, and slower average search time if they are present as distractors.

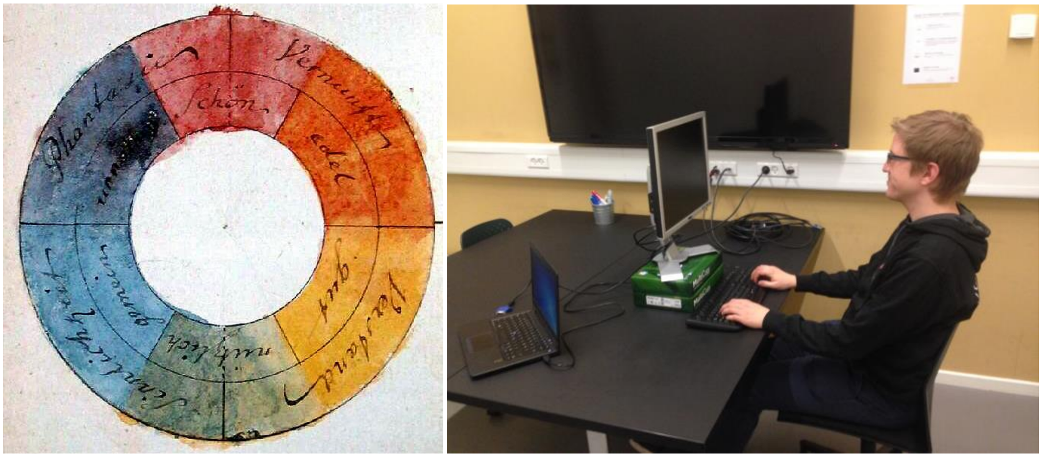


Figure 2. (left): The Colour Wheel as depicted by Johannes Wolfgang Goethe in his book Theory of Colours (1809), (right): The experiment set-up

The emotion-conveyance account states that colours convey emotions sufficiently strongly and at a sufficiently early stage in processing that it influences attentional capture. Specifically, we investigate the hypothesis that red, yellow and orange, which have generally been associated with 'action-moods' (excitement, hostility, defiance) would capture attention more than green, blue and purple, which have generally been associated with soothing or relaxing moods. As with the perception-primacy account, this would result in faster average search-times if these colours are the target, and slower average search-time if they are distractors.

3 METHODS

11 students (average age 20.9, 3 female) were recruited through an e-mail sign-up sheet distributed in person by the authors in lectures at the Technical University of Denmark. Participants were included if they were between 18 and 30 years old, had normal or corrected to normal eye-sight and did not suffer from any disorders that affect the attention system, such as ADHD.

The experiments were conducted on a Lenovo or Dell laptop computer using the E-Prime 2.0 software for Windows. The experiment was displayed on a Dell 18" monitor and participants responded via an external, USB-connected keyboard. The experiment was conducted in small quiet room with a fluorescent ceiling lamp above the participant and backlight from a glass door as the light sources. Upon arrival participants were greeted and asked to fill out a compliance form to verify that they complied with the inclusion criterion, and to give permission for the data to be used in publication. Participants were then seated ~60 cm away from the screen and with their eyes in line with the centre of the screen. Participants were furthermore equipped with Tobii Eye-Tracking glasses for the entire duration of the experiment (these results will not be reported here due to the scope of this paper). Figure 3 shows the experimental set-up.

The experiment consisted of a training block (15 trials) and experiment block (540 trials). Each trial proceeded as follows: First, a screen appeared for 1500ms, which instructed the participant which target to search for in the following task. Second, a black cross appeared in the middle of the screen for 1000ms, which the participant had been instructed to fixate on whenever it appeared. Third, a display containing 20 coloured circles appeared until the participant responded with the left keyboard arrow-key if the target was present, or the right keyboard arrow-key if the target was absent. The construction of these displays is elaborated on below. Fourth and finally, a review screen appeared which showed the participants their reaction time for the specific trial as well as their accuracy across all trials, and an instruction that the next trial could be started using the keyboard's spacebar. Figure 4 shows the experiment procedure.

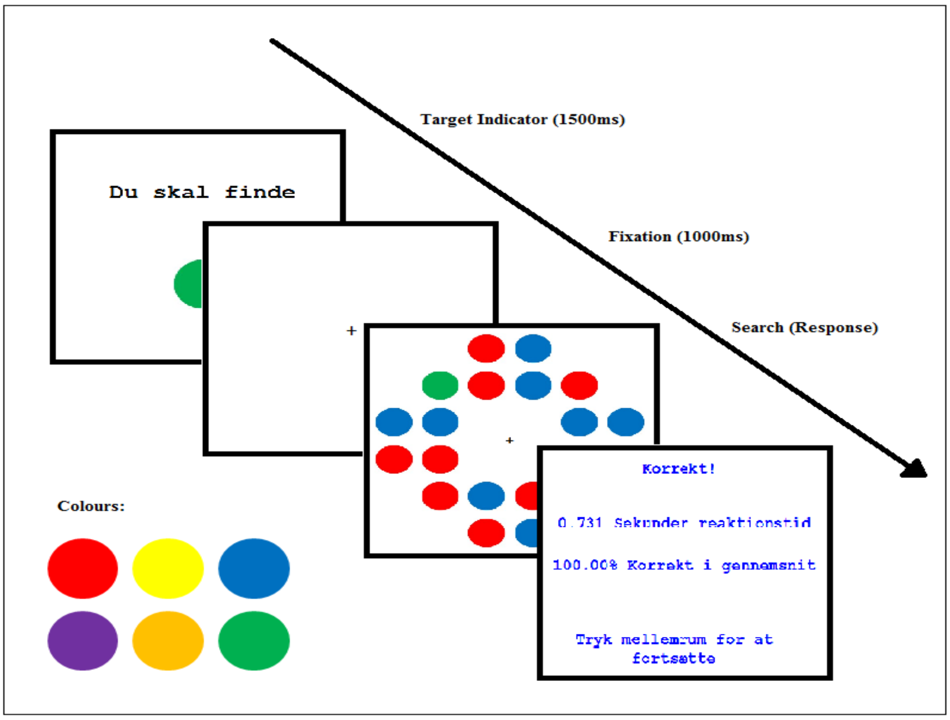


Figure 3. The experiment procedure (sample text in Danish): First, the participant is given the target colour. Second, the participant fixates on a cross in the centre of the screen. Third, the participant scans the display and responds whether the target is there or not. Fourth, the participant is given performance feedback

All stimuli displays were newly created for the experiment using MS Powerpoint and E-Prime 2.0 for Windows. A total of 360 target-present displays (2/3) and 180 target-absent displays (1/3) were created using the three primary and three secondary colours. The target-present displays were constructed each colour appeared as target with all other colours as distractors and so that the target appeared both in the inner and outer circle of the display. The distractors consisted of circles of one, two, three or four different colours and were balanced so that each of the distracting colours was approximately equally represented. The locations of both targets and specific distractor colours were randomly assigned using MS Excel's = RANDOM function. The target-absent trials were included to ensure that participants had to scan the display thoroughly before responding. They were created in the same manner as target-present displays (albeit with no target).

4 RESULTS

All analyses were conducted in the IBM SPSS statistics package v24 for Windows. The average RT was 513.73 milliseconds (ms) with a standard deviation of 156.83ms. Results were considered significant at $\alpha = 0.05$. The experiment lasted 30-35 minutes, including training, and participants showed no difference in performance as the experiment progressed. There was no significant effects of age (Spearman's $\rho = 0.25$, $p = 0.461$) or gender (Mann-Whitney $U = 17$, $p = 0.376$).

To prepare the dataset for analysis, z-values were computed for reaction time data for all responses across all participants and conditions. In total, 90 data entries were removed due to having a z-value larger than 4. As further investigation of the data showed no systematic reason for these very slow responses, it was concluded that they were due to external factors. This left a total of 14859 data entries for analysis. The analyses below are based on these data entries averaged and partitioned with respect to the participants and the specified conditions. Given that all participants had responded to all conditions, the analyses were conducted using Repeated Measures ANOVA and, if the ANOVA was significant, follow-up paired t-tests. These tests were adequate, despite the small sample size, as reaction time data was normally distributed for all parameters.

4.1 Effect of Target Colour

First, the effect of target colour on reaction time irrespective of distractor colour was analysed. A Repeated Measures ANOVA across all six target colour conditions revealed that there was a significant difference ($F_{5,50} = 18.25$, $p < 0.001$) in reaction time across the colours. Follow-up t-tests were therefore conducted to elucidate this difference. Figure 5 shows a graphical representation of the means analysed, and the results are summarized in Table 1.

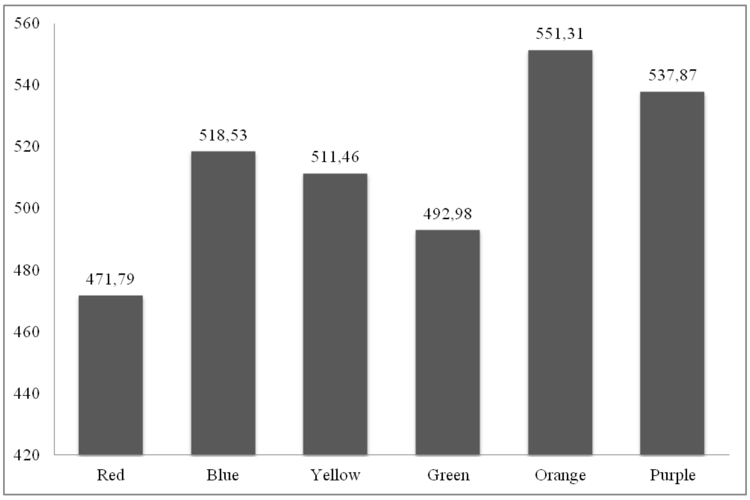


Figure 4. Reaction times (ms) for the target colours irrespective of distractor colours

Red targets were found significantly faster than blue targets (46.74 ms, $t_{10} = 5.48$, $p < 0.001$), yellow targets (39.67 ms, $t_{10} = 8.35$, $p < 0.001$), green targets (21.19 ms, $t_{10} = 6.12$, $p < 0.001$), orange targets (79.52 ms, $t_{10} = 6.55$, $p < 0.001$) and purple targets (39.77 ms, $t_{10} = 8.35$, $p < 0.001$).

Blue targets were found significantly slower than red targets (46.74 ms, t(10)=5.48, $p < 0.001$) and green targets (25.55 ms, $t_{10} = 3.02$, $p = 0.013$) and found significantly faster than orange targets (32.78ms, $t_{10} = 2.69$ $p = 0.023$) and purple targets (19.34 ms, $t_{10} = 3.25$, $p = 0.009$). There was no significant difference in reaction time between blue and yellow targets (7.07 ms, $t_{10} = 0.88$, $p = 0.402$).

Yellow targets were found significantly slower than red targets (39.77 ms, $t_{10} = 8.35$, $p < 0.001$) and green targets (18.48 ms, $t_{10} = 3.91$, $p = 0.003$) and found significantly faster than orange targets (39.85 ms, $t_{10} = 4.24$, $p = 0.002$) and purple targets (26.41 ms, $t_{10} = 2.56$, $p = 0.028$). There was no significant difference in reaction time between blue and yellow targets (7.07 ms, $t_{10} = 0.88$, $p = 0.402$).

Green targets were found significantly slower than red targets (21.19 ms, $t_{10} = 6.12$, $p < 0.001$) and significantly faster than blue targets (25.55 ms, $t_{10} = 3.02$, $p = 0.013$), yellow targets (18.48 ms, $t_{10} = 3.91$, $p = 0.003$), orange targets (58.33 ms, $t_{10} = 4.85$, $p= 0.001$) and purple targets (44.88 ms, $t_{10} = 3.75$, $p = 0.004$).

Orange targets were found significantly slower than red targets (79.52 ms, $t_{10} = 6.55$, $p < 0.001$), blue targets (32.78ms, $t_{10} = 2.69$, $p = 0.023$), yellow targets (39.85 ms, $t_{10} = 4.24$ $p = 0.002$) and green targets (58.33 ms, $t_{10} = 4.85$, $p = 0.001$). There was no significant difference in reaction time between orange and purple targets (13.45 ms, $t_{10} = 1.11$, $p = 0.29$).

Purple targets were found significantly slower than red targets (39.77 ms, $t_{10} = 8.35$, $p < 0.001$), blue targets (19.34 ms, $t_{10} = 3.25$, $p = 0.009$), yellow targets (26.41 ms, $t_{10} = 2.56$, $p = 0.028$) and green targets (44.88 ms, $t_{10} = 3.75$, $p = 0.004$). There was no significant difference in reaction time between orange and purple targets (13.45 ms, $t_{10} = 1.11$, $p = 0.29$).

Table 1. Comparison of target colour effect on reaction time (ms)

	Red	Blue	Yellow	Green	Orange	Purple
Red		-46.74**	-39.67**	-21.19 **	-79.52 **	-39.77 **
Blue	+46.74**		+7.07	+25.55*	-32.78*	-19.34**
Yellow	+39.67**	-7.07		+18.48**	-39.85**	-26.41*
Green	+21.19 **	-25.55*	-18.48**		-58.33**	-44.88 **
Orange	+79.52 **	+32.78*	+39.85**	+58.33**		+13.45
Purple	+39.77 **	+19.34**	+26.41*	+44.88 **	-13.45	

*p<0.05, **p<0.01

4.2 Effect of Distractor Colour

Second, the effect of distractor colour on reaction time irrespective of target colour was analysed. A Repeated Measures ANOVA across all six target colour conditions revealed that there was a significant difference ($F_{5,50} = 6.49$, $p < 0.001$) in reaction time across the colours. Follow-up t-tests were therefore conducted to elucidate this difference. Figure 6 shows a graphical representations of the means analysed and the analyses are summarized in Table 2.

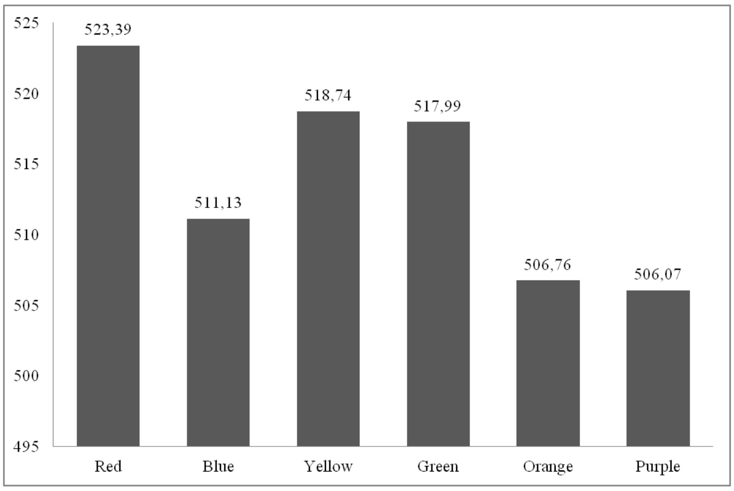


Figure 5. Reaction times (ms) for the distractor colours irrespective of target colour

Red distractors were associated with significantly slower reaction time than blue distractors (12.26 ms, $t_{10} = 3.75$, $p = 0.004$), green distractors (5.40 ms, $t_{10} = 2.65$, $p = 0.024$), orange distractors (16.62 ms, $t_{10} = 3.35$, $p = 0.007$) and purple distractors (17.32 ms, $t_{10} = 3.54$, $p = 0.005$). There was no significant difference in effect on reaction time between red and yellow distractors (4.64 ms, $t_{10} = 2.08$, $p = 0.065$).
Blue distractors were associated with significantly faster reaction time than red distractors (12.26 ms, $t_{10} = 3.75$, $p = 0.004$), yellow distractors (7.61 ms, $t_{10} = 2.39$, $p = 0.038$), and green distractors (6.86 ms, $t_{10} = 2.75$, $p = 0.020$). There was no significant difference in reaction time between blue and orange distractors (4.37 ms, $t_{10} = 0.97$, $p = 0.356$) or between blue and purple distractors (5.06 ms, $t_{10} = 1.62$, $p = 0.136$).
Yellow distractors were associated with significantly slower reaction time than blue distractors (7.61 ms, $t_{10} = 2.39$, $p = 0.038$), orange distractors (11.98 ms, $t_{10} = 2.23$, $p = 0.050$) and purple distractors (12.67 ms, $t_{10} = 2.52$, $p = 0.030$). There was no significant difference in effect on reaction time between yellow and red distractors (4.64 ms, $t_{10} = 2.08$, $p = 0.065$) or yellow and green distractors (0.75 ms, $t_{10} = 0.31$, $p = 0.762$).
Green distractors were associated with significantly slower reaction time than blue distractors (6.86 ms, $t_{10} = 2.75$, $p = 0.020$), orange distractors (11.23 ms, $t_{10} = 2.38$, $p = 0.039$) and purple distractors (11.92 ms, $t_{10} = 2.38$, $p = 0.020$) and significantly faster reaction times than red distractors (12.26 ms, $t_{10} = 3.75$, $p = 0.004$). There was no significant difference in effect on reaction time between green and yellow distractors (0.75 ms, $t_{10} = 0.31$, $p = 0.762$).
Orange distractors were associated with significantly faster reaction time than red distractors (16.62 ms, $t_{10} = 3.35$, $p = 0.007$), yellow distractors (11.98 ms, $t_{10} = 2.23$, $p = 0.050$) and green distractors (11.23 ms, $t_{10} = 2.38$, $p = 0.039$). There was no significant difference in effect on reaction time between orange and blue distractors (4.37 ms, $t_{10} = 0.97$ $p = 0.356$) or between orange and purple distractors (0.69 ms, $t_{10} = 0.18$, $p = 0.863$).
Purple distractors were associated with significantly faster reaction time than red distractors (17.32 ms, $t_{10} = 3.54$, $p = 0.005$), yellow distractors (12.67 ms, $t_{10} = 2.52$, $p = 0.030$) and green distractors (11.92 ms, $t_{10} = 2.38$, $p = 0.020$). There was no significant difference in effect on reaction time between purple and blue distractors (5.06 ms, $t_{10} = 1.62$, $p = 0.136$) or between purple and orange distractors (0.69 ms, $t_{10} = 0.18$, $p = 0.863$).

Table 2. Comparison of distractor colour effect on reaction time

	Red	Blue	Yellow	Green	Orange	Purple
Red		+12.26 **	+4.64	+5.40*	+16.62**	+17.32 **
Blue	-12.26 **		-7.61*	-6.86 *	+4.37	+5.06
Yellow	-4.64	+7.61*		+0.75	+11.98*	+12.67*
Green	-5.40*	+6.86 *	-0.75		+11.23*	+11.92*
Orange	-16.62**	-4.37	-11.98*	-11.23*		+0.69
Purple	-17.32 **	-5.06	-12.67*	-11.92*	-0.69	

*p<0.05, **p<0.01

5 DISCUSSION

In this paper we presented a study of the attentional capture of red, blue, yellow, green, orange and purple. In accordance with our main hypothesis, different colours were associated with different degrees of attentional capture, with red capturing attention significantly more than all other colours and purple and orange capturing significantly less attention than the remaining. To explain why different colours capture attention differently, we proposed and investigated what we referred to as the perceptual-primacy account and the emotion-conveyance account. In the following sections, we discuss the evidence for and against each of these two accounts, as well as potential implications for the disciplines of design and psychology.

5.1 The Perceptual-Primacy Account

The perceptual-primacy account states that colours denoted as primary colours receive preferential processing and thereby capture attention to a larger extent. We here investigated two versions of this account: one which denotes the primacy of colours based on our retinas, which posits that red, green and blue are the primary colours, and one which denotes the primary colours based on the properties of pigments, which posits that the primary colours are red, blue and yellow.

Our results show support for both versions of the perceptual-primacy account: Red and green targets were found significantly faster than blue and yellow targets, which in turn were found significantly faster than orange and purple targets. In accordance with the notion of higher attentional capture of these colours, it was further found that, with one exception, average search times were slowed significantly more when a red, blue, yellow or green distractor was present.

However, there were significant differences in search times amongst the primary colours when compared both on their role as a target and a distractor. Of these, two were especially notable: One, with a single exception, red always captured significantly more attention than all other colours. Two, yellow performed differently compared to the other colours as a target and a distractor; when comparing targets, yellow captured significantly more attention than green, and did not differ from blue, when comparing distractors, yellow captured significantly more attention than green and was not significantly different from red in its effect. While these findings are not in disagreement with the perceptual-primacy account, they do suggest that other factors influence differences in attentional capture from colour as well. We discuss these alternate accounts below, beginning with the emotional-conveyance account.

5.2 The Emotional-Conveyance Account

The emotional-conveyance account states that colours convey emotions to a sufficiently salient degree and at a sufficiently early stage of processing to influence attentional capture. We here investigated this account by comparing colours associated with ‘action moods’ with ‘soothing and relaxing moods’. This assumed that if colours convey emotions at an early stage then the colours associated with particularly salient colours would receive preferential processing.

Our results showed evidence against the emotional-conveyance account, albeit inconclusively. In support of the emotional-conveyance account, red, which, as stated in the introduction, is associated with exciting and defying/hostile moods, captured significantly more attention than any other colours as both a target and all colours but yellow a distractor, which was not significantly different to red. Furthermore, yellow, which is associated with exciting moods, captured the same amount of attention as red, and more attention than any other colour as a distractor. However, yellow targets were not found significantly faster than red, blue or green and orange, which is also associated with exciting moods, consistently captured the least amount of attention alongside purple. The results thus indicate that exciting and relaxing moods were not conveyed at a sufficiently early stage to influence attentional capture. Furthermore, given that red did not always capture attention more than other colours, it seems unlikely that the emotion-conveyance account explains the observed differences.

5.3 Alternate Accounts

Given that neither the perceptual-primacy account or emotional-conveyance account were completely congruent with the data, we here discuss two potential alternate influencing factors.

First, it is possible that training effects and conventional uses of the colours have influenced the search tasks. For example, red is commonly used in traffic and to signal important information for emergencies. Participants may therefore be trained in searching for red, which would explain why red showed significantly more attentional capture as a target, but was not significantly different from yellow in attentional capture as a distractor. Second, it may be that we are inherently predisposed towards certain colours due to other biological factors than the ones considered here, such as gathering of certain foods or similar.

5.4 Insights for Designers

Our results showed colour mediated differences in reaction time in a visual search task. Furthermore, we found that it is unlikely that colours convey moods and emotion until conscious processing has occurred.

While the differences were on the scale of tens to hundreds of milliseconds, these effects may be amplified in contexts of higher visual load (Lavie and Tsal, 2004), or simply add up over time. We therefore propose that visual designs should use red or green for objects that are particularly important for the user to find quickly, and to avoid using red and yellow for miscellaneous objects in an interface, as they would interfere with visual search. Furthermore, our results suggest that designers will not be successful in rapidly conveying emotion or mood tones through colour.

5.5 Insights for Psychology

Studies of the degree to which colour can capture attention have thus far shown diverging results (Folk et al., 1994; Theeuwes, 1994; Müller et al., 2009). In the light of our finding that different colours have different attentional effects, we suggest that future studies should either replicate the choice of colours used by their predecessors, or, preferably, to investigate attentional effects for an array of different colours, given that the results might diverge in accordance with the colours used in the experiment design.

On a more general note, our results imply that great care should be taken in experiment designs where colours are used to indicate specific objectives for the experiment. For example, when conducting a partial report task, one may for example be asked to report all red letters and to ignore blue letters. Our results indicate that these types of tasks could be biased simply by the colours used, rather than any other properties one may have wanted to investigate. To avoid this, careful choice of colours and/or counterbalancing would be necessary.

6 CONCLUSION

In this paper, we presented a controlled-environment study of the attention-capturing properties of red, blue, yellow, green, orange, purple. Specifically, we tested whether effects could be explained through a primary-secondary colour dichotomy and/or through differences in emotional conveyance of colours. We found significant differences in attentional capture between primary and secondary colours, defined both from a biological and design perspective. Notably, red captured significantly more attention than other colours, while orange and purple consistently captured the least attention. However, we did not find substantial evidence to support that colours convey emotions at an early enough stage to affect attentional capture. Alternate accounts and impacts for design and psychology were discussed.

REFERENCES

Braseth, A.O., Øritsland, T.A. (2013), Visualizing complex processes on large screen displays: Design principles based on the Information Rich Design concept. *Displays,* Vol. 34, pp. 215–222.

Brosch, T., Sharma, D. (2005), The Role of Fear-Relevant Stimuli in Visual Search: A Comparison of Phylogenetic and Ontogenetic Stimuli. *Emotion*, Vol. 5, pp. 360–364. doi:10.1037/1528-3542.5.3.360

Desmet, P.M. (2015), Design for mood: Twenty activity-based opportunities to design for mood regulation. *International Journal of Design*, Vol. 9 (2), 2015.

Folk, C.L., Remington, R.W., Wright, J.H. (1994), The structure of attentional control: contingent attentional capture by apparent motion, abrupt onset, and color. *Journal of Experimental Psychology: Human perception and performance*, Vol. 20, pp. 317.

Gazzaniga, M.S. (2009), Cognitive neuroscience, the biology of the mind, 3. ed. ed. WWNorton.

Gillet, R. (2014), What Your Logo's Color Says About Your Company (Infographic) [WWW Document]. Fast Company. URL https://www.fastcompany.com/3028378/leadership-now/what-your-logos-color-says-about-your-company-infographic (accessed 12.19.16).

Harkness, N. (2006), The colour wheels of art, perception, science and physiology. Optics *& Laser Technology*, Vol. 38, pp. 219–229. https://doi.org/10.1016/j.optlastec.2005.06.10.

Hodsoll, S., Viding, E., Lavie, N. (2011), Attentional capture by irrelevant emotional distractor faces. *Emotion*, Vol. 11, pp. 346–353. doi:10.1037/a0022771

Huang, Y.-M., Baddeley, A., Young, A.W. (2008), Attentional capture by emotional stimuli is modulated by semantic processing. *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 34, pp. 328–339. doi:10.1037/0096-1523.34.2.328

Koster, E.H.W., Crombez, G., Van Damme, S., Verschuere, B., De Houwer, J. (2004), Does Imminent Threat Capture and Hold Attention? *Emotion*, Vol. 4, pp. 312–317. doi:10.1037/1528-3542.4.3.312

Lau, N., Veland, Ø., Kwok, J., Jamieson, G.A., Burns, C.M., Braseth, A.O., Welch, R. (2008), Ecological Interface Design in the Nuclear Domain: An Application to the Secondary Subsystems of a Boiling Water Reactor Plant Simulator. *IEEE Transactions on Nuclear Science*, Vol. 55, pp. 3579–3596. doi:10.1109/TNS.2008.2005979

Lavie, N., Hirst, A., de Fockert, J.W., Viding, E. (2004), Load Theory of Selective Attention and Cognitive Control. *Journal of Experimental Psychology: General*, Vol. 133, pp. 339–354. doi:10.1037/0096-3445.133.3.339

Lee, J., Jung, B., Chu, W. (2015), Signaling environmental altruism through design: The role of green cue prominence in hybrid cars. *International Journal of Design*, Vol. 9, pp. 79–91.

Müller, H.J., Geyer, T., Zehetleitner, M., Krummenacher, J. (2009), Attentional capture by salient color singleton distractors is modulated by top-down dimensional set. *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 35, pp. 1. doi:https://doi.org/10.1016/j.displa.2013.05.002

Murray, D.C., Deabler, H.L. (1957), Colors and mood-tones. *Journal of Applied Psychology*, Vol. 41, pp. 279. doi:https://doi.org/10.1037/h0041425

Na, N., Suk, H.-J. (2014), The Emotional Characteristics of White for Applications of Product Color Design. *International Journal of Design*, Vol. 8, pp. 61–70. doi:https://doi.org/10.1007/978-3-319-04798-0_20

Ohman, A., Flykt, A., Esteves, F. (2001), Emotion drives attention: Detecting the snake in the grass. *Journal Of Experimental Psychology-General*, Vol.130, pp. 466–478. doi:https://doi.org/10.1037//0096-3445.130.3.466

Page, T., Thorsteinsson, G., Ha, J.-G. (2012), Using colours to alter consumer behaviour and product success. *International Journal of Contents*, Vol. 8, pp. 69–73. doi:https://doi.org/10.5392/ijoc.2012.8.1.069

Schaie, K.W. (1961a), A Q-sort study of color-mood association. *Journal of Projective Techniques*, Vol. 25, pp. 341–346. doi:https://doi.org/10.1080/08853126.1961.10381048

Schaie, K.W. (1961b), Scaling the association between colors and mood-tones*. The American Journal of Psychology*, Vol. 74, pp. 266–273. doi:https://doi.org/10.2307/1419412

Theeuwes, J. (1994), Stimulus-driven capture and attentional set: selective search for color and visual abrupt onsets. *Journal of Experimental Psychology: Human perception and performance*, Vol. 20, pp. 799. doi:https://doi.org/10.1037//0096-1523.20.4.799

Vazquez, E., Gevers, T., Lucassen, M., Van de Weijer, J., Baldrich, R. (2010), Saliency of color image derivatives: a comparison between computational models and human perception. *JOSA A*, Vol. 27, pp. 613–621. https://doi.org/10.1364/JOSAA.27.000613

Ware, C. (2010), Visual thinking: For design. Morgan Kaufmann.

Wexner, L.B. (1954), The degree to which colors (hues) are associated with mood-tones. *Journal of Applied Psychology*, Vol. 38, pp. 432. https://doi.org/10.1037/h0062181

Paper 4:

The Attentional Guidance of Colour in Increasingly Complex Displays (submitted),

by Andersen, E., & Maier, A.

Submitted to Applied Ergonomics

The Attentional Guidance of Individual Colours in Increasingly Complex Displays

Emil Andersen* and Anja Maier

Technical University of Denmark

Department of Management Engineering, Engineering Systems Division,

Produktionstorvet 424, Kgs. Lyngby, Denmark

*Corresponding author e-mail address: emand@dtu.dk

Abstract

The use of colors is a prevalent and effective tool for improving design. Understanding the effect of colors on attention is crucial for designers that wish to understand how their interfaces will be used. Previous research has consistently shown that attention is biased towards color. However, despite previous evidence indicating that colors should be treated individually, it has thus far not been investigated whether this difference is reflected in individual effects on attention. To address this, a visual search experiment was conducted that tested the attentional guidance of six individual colors (red, blue, green, yellow, orange, purple) in increasingly complex displays. Results showed significant that the individual colors differed significantly in their level of guidance of attention, and that these differences increased as the visual complexity of the display increased. Implications for visual design and future research on applying color in visual attention research and design are discussed.

1. Introduction

As computer technology has progressed and become more popular, the prevalence and importance of graphical interfaces, too, has increased. This tendency includes computer interfaces for use in situations, such as medicine, military and nuclear power plants, where improper action may have dramatic consequences. For example, improper design of the software used by Hawaii Emergency Management Agency has been identified as possibly contributing to an operator falsely transmitting a warning about an impending missile strike to the Hawaiian population (Federal Communicatinos Comission, 2018). A common design requirement for proper interface design is that users must be able to find specific, task-relevant objects amongst task-irrelevant objects. To achieve this, designers commonly and effectively (Vazquez, Gevers, Lucassen, van de Weijer, & Baldrich, 2010) use colour to separate and group objects in visual design. For example, Starke & Barber (2018) showed that user interfaces with colour lead to faster (and equally accurate) information foraging in a complex credit card fraud detection task. Similarly, previous research on colour in applied contexts has informed guidelines for ergonomic visual design with regards to how to match object size and colour (Poulton, 1975), improve legibility through colour use for text presented on CRT (ISO 9241-306:2018; Matthews, 1987) and LCD screens (Humar, Gradisar, Turk, & Erjavec, 2014), optimize maps through correct colour use (Francis, Bias, & Shive, 2010; Shive & Francis, 2013), and for optimizing colour background-icon combinations (Huang, Baddeley, & Young, 2008).

The importance of colour in visual design is corroborated by its importance in experimental psychology, where colour has been shown to reliably bias our attention, meaning that search is facilitated for coloured targets, and that attention is more easily captured by coloured distractors (Biggs, Kreager, & Davoli, 2015; Eimer & Grubert, 2014; Eimer & Kiss, 2010; Folk, Remington, & Wright, 1994; Grubert & Eimer, 2015; Muhl-Richardson et al., 2018; Müller, Geyer, Zehetleitner, & Krummenacher, 2009; Nordfang, Dyrholm, & Bundesen, 2013; Theeuwes, 1994; Wolfe & Horowitz, 2017, 2004). Notably, whereas the aforementioned applied studies mostly consider how visual designs of various colour combinations perform, the experimental studies of colour and attention mostly consider colour as a group (i.e. does 'colour' capture attention). The resulting research has thus aided our knowledge on combinations of colour and the attentional effects of colour as a group. However, little research has been performed on the relation between attention and specific colours. In this paper, we expand our previous discussion and findings (Andersen & Maier, 2017) on why individual colours should be expected to have distinct properties in attentional guidance and capture. We test them using an experimental design that combines the rigorous methods of experimental psychologists with the expected complexity of a realistic visual design problem.

2. Background

2.1 Visual Design and Complex Displays

As argued by Zacks & Tversky (2003), the use of top-down principles can be greatly beneficial for design of visual interfaces. By applying knowledge acquired through studies of human behaviour to design, predictions can be made that provide insights for how to better (or optimally) create visual interfaces. Examples of such insights range from websites for everyday usability (Tsai, Chang, Chuang, & Wang, 2008), maps (Francis et al., 2010; Shive & Francis, 2013), graphs (Ratwani, Trafton, & Boehm-Davis, 2008) or advertising and branding (Ko, 2017; Page, Thorsteinnson, & Ha, 2012; Resnick & Albert, 2016), to the design of nuclear control room interfaces (Braseth & Øritsland, 2013; Lau et al., 2008; Van Laar & Deshe, 2002, 2007) and hazard information systems (Miran, Ling, James, Gerard, & Rothfusz, 2017). In the same vein, design guidelines based on insights from attention theory has shown promise for improving performance of users of complex visual interfaces (McCarley & Steelman, 2013): Understanding the attentional system is particularly important for the design of complex visual interfaces that contain large amounts of information, as search for the task-relevant objects can be complicated by the limited processing capacity of our attention system. As the foundational study by Treisman & Gelade (1980) showed, search in these situations depends on the similarity of the objects: If objects are similar, only a few items can be considered at a given moment, resulting in slow, serial search. However, if the object is distinguished from other objects by a feature that can guide attentional allocation, search will be facilitated by a pre-attentive evaluation of the objects (for later corroborating accounts see e.g. Bundesen, 1990; Bundesen, Habekost, & Kyllingsbæk, 2005; Wolfe, Cave, & Franzel, 1989; Wolfe & Gray, 2007). Therefore, knowing which features guide and capture attention allows designers to create visual interfaces that are easy, pleasant and intuitive to use, through e.g. optimal use of colour coding (Travis, 1990; Travis, Bowles, Seton, & Peppe, 1990), colour highlighting (McDougald & Wogalter, 2014) or visual layering (Van Laar & Deshe, 2002, 2007). As outlined above, an often-used and well-established guiding factor has been found in colour (for a list of other potential candidates see Wolfe & Horowitz, 2004, 2017).

2.2 Attentional Guidance of Colour

A large body of work has shown that colour as a feature will capture attention if it is a distractor (e.g. Snowden, 2002; Theeuwes, 1992, 1994) or guide attention if it is a target (Biggs et al., 2015;

Eimer & Grubert, 2014; Eimer & Kiss, 2010; Folk et al., 1994; Grubert & Eimer, 2015; Müller et al., 2009; Nordfang et al., 2013; Wolfe, 2007; Wolfe et al., 1989; Wolfe & Horowitz, 2004, 2017). For example, Theeuwes (1992) asked participants to search for lines embedded in a green circle amongst green squares, or green squares with a single-ton red square. He found that the single-ton red distractor slowed reaction time, thereby showing that attention had been captured by the distracting colour object. This finding was corroborated by Nordfang and colleagues (2013), who further found that the attentional capture of a colour singleton is amplified if it is a target. Recently, Biggs and colleagues (2015) expanded this finding, showing that the guiding of colour depends on the visual load of a display. By integrating the aforementioned Guided Search account with Load Theory (Lavie, 2010; Lavie, Hirst, De Fockert, & Viding, 2004; Lavie & Tsal, 1994), Biggs and colleagues (2015) showed found that colour only guided search under high visual load. However, while these studies provide evidence that colour guides and captures attention (to a varying degree, depending on the context), they have in common that they treat colour as a group, rather than investigating the effect on attention of specific colours. However, in applied contexts, colours are seldom treated as a group. Instead specific individual colours are used for to for example comply with style choices or colour contrast guidelines. If the attentional effects of colours differ for specific colours, colours should be considered individually for design to avoid unintended bias in navigation. While international standards and recommendations exist for measuring and optimizing colour usage with regards to e.g. legibility and fidelity (ISO 9241-306:2018, 2018; Travis, 1990; Travis et al., 1990), to our knowledge, no such recommendations have so far been made with regards to attentional capture and guidance of specific colours. As will be discussed below, however, there are several reasons why the specific colours may affect attention differently.

2.3 Reasons for Individual Effects of Colour

The retinas of our eyes contain a myriad of two types of photoreceptive cells commonly referred to as rods and cones. Of these, rods are sensitive to white and black, whereas cones are primarily sensitive to one of red, blue or green. All other colours are made through combinations of responses from these cells (Gazzaniga, Ivry, & Mangun, 2009; Ware, 2008). Individual colours are thus treated separately at the cellular level, even before the colour signal reaches consciousness. In addition the separation of colours into primary and secondary based on the properties of pigments (Harkness, 2006), is widely distributed and used in design and the arts. Given these micro and macro separations of colours, it could be expected that individual colours are treated differently as well. Indeed, individual colours have been shown extensively to be associated with individual emotional values in subjective rating experiments: Red is most frequently associated with exciting, defending and defying/hostile moods, blue with soothing and secure, yellow and orange with stimulating, cheerful and exciting moods, green with calm and soothing and purple with dignified and stately moods (Aslam, 2006; Lee, Jung, & Chu, 2015; Murray & Deabler, 1957; Na & Suk, 2014; Schaie, 1961b, 1961a; Wexner, 1954). Given that previous research has shown that emotional stimuli are related with higher attentional priority (Brosch & Sharma, 2005; Hodsoll, Viding, & Lavie, 2011; Huang et al., 2008; Koster, Crombez, Van Damme, Verschuere, & De Houwer, 2004; Ohman, Flykt, & Esteves, 2001; Öhman, Lundqvist, & Esteves, 2001) it could be extended that individual colours may have different effects on the attention system based on their emotional valence. Based on these findings, the present paper tested whether these differences between colours were reflected in different effects on attentional guidance and capture.

2.4 Hypotheses: Testing the Effect of Individual Colour on Attentional Guidance and Capture

Using a visual search experiment wherein participants searched for coloured targets amongst other colored targets, we tested four hypotheses:

- **Hypothesis 1:** Specific colours will guide and capture attention differently.

Given the above findings on the differences between individual colours, this hypothesis can be specified as

- Hypothesis 1a: Specific colours will guide and capture attention differently in concordance with the primary – secondary divide found in the eye or the properties of pigments.
- Hypothesis 1b: Specific colours will guide and capture attention due to their emotional properties.

In order to investigate the primary-secondary hierarchies, the colours used were red, green, blue, yellow, orange and purple. Furthermore, based on the findings of Biggs and colleagues (2015), we tested the hypothesis 2:

- **Hypothesis 2:** The effect of individual colours on attention increases as the complexity (or visual load) increases.

3. Methods

3.1 Participants

17 students were recruited through physical e-mail sign-up sheets. Participants were included if they were between 18 and 30 years old, had normal or corrected to normal eye-sight and did not suffer from any disorders that affect the attention system, such as ADHD.

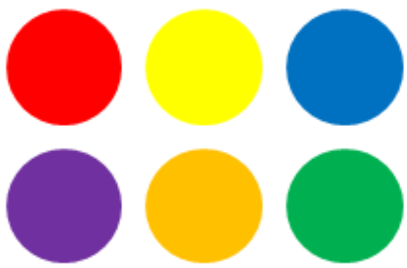
3.2 Apparatus

The experiments were conducted on a laptop computer using the E-Prime 2.0 software for Windows. The experiment was displayed on an 18" monitor and participants responded via an externally connected keyboard. The experiment was conducted in small, well-lit room. Upon arrival participants were greeted and asked to fill out a compliance form to verify that they complied with the inclusion criterion, and to give permission for the data to be used in publication. Participants were seated ~60 cm away from the screen and with their eyes in line with the centre of the screen.

3.4 Stimuli

All stimuli displays were newly created for the experiment using MS PowerPoint and E-Prime 2.0 for Windows. A total of 360 target-present displays (2/3) and 180 target-absent displays (1/3) were created using six colours: Red, blue, green, yellow, orange and purple. The colour codes, contrast scores and luminance measures for each colour are shown in table 1. The colours are shown in figure 1. Each display contained 20 coloured circles measuring 1.9˚ visual angle in diameter. The target-present displays were constructed so that each colour appeared as a target with all other colour combinations as distractors and so that the target appeared both in the inner and outer circle of the display. The distractors consisted of circles of one, two, three or four different colours and were balanced so that each of the distracting colours was approximately equally represented. The locations of both targets and specific distractor colours were randomly assigned using MS Excel's = RANDOM function. Target-absent trials were included to ensure that participants had to scan the display thoroughly before responding. They were created in the same manner as target-present displays (albeit with no target).

Colour	R	G	B	Contrast	Luminance
Red	255	0	0	4	29
Yellow	255	255	0	1.1	106
Green	0	176	80	2.9	33
Blue	0	112	192	5.1	15
Orange	255	192	0	1.6	66
Purple	112	48	160	8	7



On the left, Table 1: Colour codes, contrast scores and luminance scores (cd/m^2), on the right Figure 1: The colours used in the experiment. See the online article for a coloured version of this figure.

3.5 Procedure

The experiment consisted of a training block (15 trials) and experiment block (540 trials). Each trial proceeded as follows: First, a screen appeared for 1500ms, which instructed the participant which target to search for in the following task. Second, a fixation cross appeared in the middle of the screen for 1000ms. Third, the search display appeared until the participant responded with the left keyboard arrow-key if the target was present, or the right keyboard arrow-key if the target was absent. Fourth and finally, a review screen appeared which showed the participants their reaction time for the specific trial as well as their accuracy across all trials, and an instruction that the next trial could be started using the keyboard's spacebar. Figure 2 shows the experiment procedure.

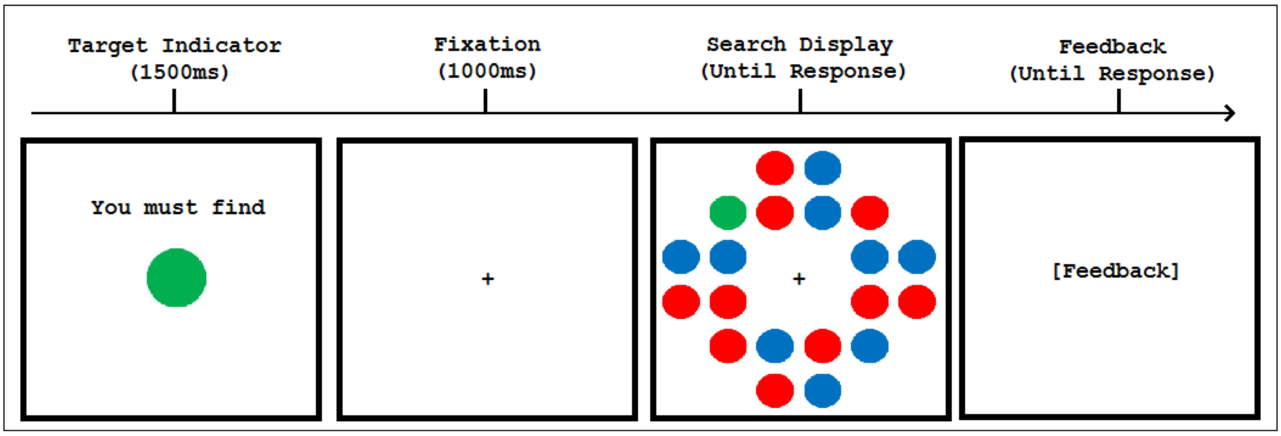


Figure 2. The experiment procedure. See the online article for a coloured version of this figure.

3.6. Analysis

Target-absent trials were excluded from the dataset prior to analysis. Differences in target colour were analysed using a 2x4 factorial repeated measures ANOVA model. The dependent variable was Reaction Time (RT). The independent variables were target or distractor colour (6 levels) and number of distractor colours (4 levels, 1 to 4 distractor colours). Subject was included as a random factor to remove inter-subject variability. Gender was tested as a co-variate given previous associations with attention control and colour perception (Harter, Miller, Price, LaLonde, & Keyes, 1989; Naglieri & Rojahn, 2001), but removed from the final analysis as it did not significantly predict Reaction Time. While using parametric tests can sometimes be problematic in studies using smaller sample sizes, the assumptions of the ANOVA model were fulfilled for the present experiment after log transforming the dependent variable, RT, and as such analysis could proceed. Means and figure data show untransformed data. Analysis was conducted using base R and the nlme package for R.

4. Results

4.1 Difference Between Individual Target Colours

There was a significant difference between individual target colours ($F = 51.27$, $p < 0.001$) and with number of distractor colours ($F = 32.45$, $p < 0.001$). There was a significant interaction effect ($F = 2.39$, $p = 0.002$), indicating that the difference in search speed for individual target colours increased as the number of distractor colours increased (see figure 3).

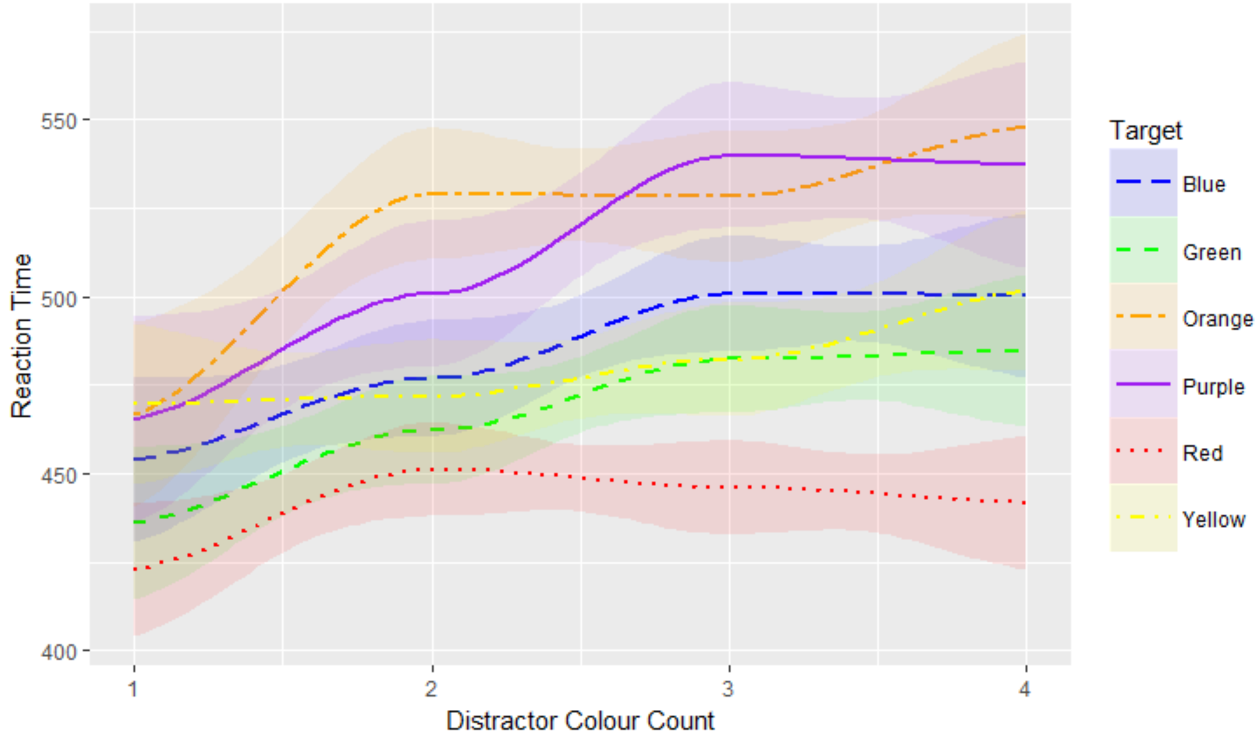


Figure 3: The interaction between target colour and amount of different distractor colours. Error bars indicate 95% confidence intervals. See the online article for a coloured version of this graph.

4.2 Difference Between Individual Distractor Colours

There was a significant difference between individual distractor colours ($F = 5.64$, $p < 0.001$) and with number of distractor colours ($F = 44.63$, $p < 0.001$). There was no significant interaction effect ($F = 1.28$, $p = 0.20$), indicating that the difference in search speed for individual distractor colours did not increase as the number of distractor colours increased. Post-hoc tests revealed that red slowed search significantly more than blue ($p = 0.016$), orange ($p < 0.001$) and purple ($p < 0.001$), while yellow slowed search significantly more than orange ($p = 0.025$) and purple ($p < 0.01$). Finally, green slowed search significantly more than purple ($p = 0.019$).

5. Discussion

This paper investigated the hypothesis that specific colours would guide and capture attention differently. Furthermore, based on findings by Biggs et al., (2015), the hypothesis was tested that any attentional guidance would increase as the complexity of the display increased. Our results support the hypothesis of a difference between individual colours, showing a three-way grouping: 1) red is consistently found the fastest, 2) blue, green and yellow are found slower than red, but 3) faster than orange and purple, which are found equally fast. Furthermore, our results support the hypothesis that differences in search times between individual colours increase as the complexity of the screen, induced through a higher number of differently coloured distractors, is increased. Finally, our data shows that the differences in attentional guidance are also reflected in increased attentional capture of the colours, with red, yellow and green slowing search significantly when present.

As discussed in the introduction, several reasons could cause individual differences between colours. Our results were consistent with hypotheses 1a and 1b for the colour red: red is specialized for on the cellular level, is considered a primary colour in design and art conventions and is frequently associated with threat and danger. However, for the remaining colours, neither hypothesis 1a or 1b fit the data: Explaining the data based on the artistic convention (i.e. red-blue-yellow is considered primary and green-orange-purple is considered secondary), is inconsistent with green being found equally fast as yellow and blue, and with red being found the fastest. Explaining based on the biology of our eyes is inconsistent with yellow being found equally fast as blue and green, and, again, is inconsistent with red being found the fastest. Explaining based on the emotional value of the colours is inconsistent with regards to how colours are grouped, as colours that are similar in their emotional meaning do not capture the same amount of attention (e.g. orange and yellow). Furthermore, the results show that red was particularly efficient at facilitating search, to the extent that search for red objects was not affected by increased complexity. It thus seemed that, unlike other colours, participants were always able to perform a parallel search for the colour red at all levels of complexity. Instead, our results show a three-group structure as the best fit for our data: 1) red, 2) green, blue and yellow, and 3) purple and orange. While neither proposed cause was supported the data, our results nevertheless indicate that specific colours have different effects on attentional guidance and capture, and that this effect increases as complexity of the display increases. Based on this finding, implications for attention theory and visual design are discussed.

5.1. Implications for Attention Studies on Colour

Our results may contribute to both the theoretical understanding of guidance by colour, which serves as a part of several separate accounts of attention (e.g. Bundesen, 1990; Wolfe et al., 1989), and our understanding of how colour conveys emotion.

First, while previous studies have mostly attributed the guiding and capturing effect of colour to the difference in contrast (e.g. Nordfang et al., 2013, Theeuwes, 1992;1994), our results do not corroborate this account. For example, red, which guided attention the best had a medium contrast (4), while the two colours that guided attention the least, purple and orange, had the highest (8) and second-lowest (1.6) salience scores respectively. Our results thus corroborate those of Snowden (2002) in showing that colour can capture attention in itself in a manner that is separate from luminance or contrast.

Second, while colours have previously been established as being associated with certain emotions (Lee et al., 2015; Murray & Deabler, 1957; Na & Suk, 2014; Schaie, 1961b, 1961a; Wexner, 1954), our results indicate that this emotional conveyance is either not sufficiently strong for it to guide attention, or does not occur until a later stage of processing, rendering it unable to influence attentional allocation. Of these two accounts, we find the latter explanation more likely, given the success in reproducing colour – emotion associations across diverse groups (Aslam, 2006). However, further research is needed to elucidate this relationship.

5.2 Implications for Visual Design

As discussed in the introduction, insights on human behaviour can inform the design of visual interfaces, which in turn may lead to improved performance of their users. Indeed, the results presented here may have several applications in the design of complex visual interfaces:

First, our findings could be applied to create colour schemes for objects so that their colour is aligned with their importance. Given our results, the highest priority object could thus be coloured red, whereas the least important object could be coloured purple. Our findings further show that this alignment is particularly important for the design of complex visual interfaces, as the differences in guidance by colour increased with display complexity. As introduced previously, such alignment of importance and colour of objects has previously been shown to improve graph (Ratwani et al., 2008) and map design (Francis et al., 2010; Shive & Francis, 2013). Our findings add to these by giving a general hierarchy of the role of specific colours on attention allocation, which may be useful for the design of complex visual displays, or e.g. for optimizing colour highlighting techniques such as those reported by McDougald & Wogalter (2014). Furthermore, in compliment to previous studies showing the trade-off of benefits between colour variation and luminance variation (Travis et al., 1990), the results show difference in attentional guidance and capture that could not be explained by differences in contrast or luminance.

Second, our finding that it may require elaborate processing before colours convey emotion can have significant consequences for how colours are used to convey semantic knowledge about products over short durations, such as in advertising, where colour plays a central role for signalling specific traits such as environmental friendliness (Aslam, 2006; Lee et al., 2015).

6. Conclusion

Colours as a group have been consistently shown to guide and capture attention. This paper investigated experimentally whether individual colours have different levels of guidance, and whether this guidance is contingent on the complexity of the display. The results supported both these claims, showing significant differences between colours that became larger as a function of interface complexity. The data suggested a three-group structure, with 1) red facilitating search the most, 2) green, blue and yellow the second-most, and 3) orange and purple facilitating search the least. Implications for theory of perception and application to design were discussed.

7. Acknowledgments

Part of the data for this study served as the basis for a contribution to the 21st International Conference on Engineering Design (ICED17). We thank the study participants at the Technical University of Denmark for their participation, and our research partners from the Halden HAMMLAB at the Institute for Energy Technology (IFE) in Norway for their continued support. The work reported in this paper is funded by The Technical University of Denmark and the OECD Halden Reactor Project.

8. Literature

Andersen, E., & Maier, A. (2017). The Attentional Capture of Colour in Visual Interface Design : A Controlled-Environment Study. In *Proceedings of the International Conference on Engineering Design, ICED* (Vol. 8, pp. 519–528).

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. https://doi.org/10.1080/13527260500247827

Biggs, A. T., Kreager, R. D., & Davoli, C. C. (2015). Finding a link between guided search and perceptual load theory. *Journal of Cognitive Psychology*, *27*(2), 164–179. https://doi.org/10.1080/20445911.2014.987676

Braseth, A. O., & Øritsland, T. A. (2013). Visualizing complex processes on large screen displays: Design principles based on the Information Rich Design concept. *Displays*, *34*(3), 215–222. https://doi.org/10.1016/j.displa.2013.05.002

Brosch, T., & Sharma, D. (2005). The role of fear-relevant stimuli in visual search: A comparison of phylogenetic and ontogenetic stimuli. *Emotion*, *5*(3), 360–364. https://doi.org/10.1037/1528-3542.5.3.360

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, *97*(4), 523–547. https://doi.org/10.1037/0033-295X.97.4.523

Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2005). A neural theory of visual attention: Bridging cognition and neurophysiology. *Psychological Review*, *112*(2), 291–328. https://doi.org/10.1037/0033-295X.112.2.291

Eimer, M., & Grubert, A. (2014). The gradual emergence of spatially selective target processing in visual search: {From} feature-specific to object-based attentional control. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(5), 1819–1831. https://doi.org/10.1037/a0037387

Eimer, M., & Kiss, M. (2010). Top-down search strategies determine attentional capture in visual search: Behavioral and electrophysiological evidence. *Attention, Perception, and Psychophysics*, *72*(4), 951–962. https://doi.org/10.3758/APP.72.4.951

Federal Communicatinos Comission. (2018). Presentation on Preliminary Report on Hawaii False Emergency Alert. *Federal Communications Commission*. Retrieved from https://www.fcc.gov/document/presentation-preliminary-report-hawaii-false-emergency-alert

Folk, C. L., Remington, R. W., & Wright, J. H. (1994). The structure of attentional control: Contingent attentional capture by apparent motion, abrupt onset, and color. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(2), 317–329. https://doi.org/10.1037/0096-1523.20.2.317

Francis, G., Bias, K., & Shive, J. (2010). The Psychological Four-Color Mapping Problem. *Journal of Experimental Psychology: Applied*, *16*(2), 109–123. https://doi.org/10.1037/a0019095

Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (2009). *Cognitive Neuroscience: The Biology of the Mind*. *The Quarterly Review of Biology* (3. ed., Vol. 84). WWNorton. https://doi.org/10.1086/603482

Grubert, A., & Eimer, M. (2015). Rapid parallel attentional target selection in single-color and multiple-color visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(1), 86–101. https://doi.org/10.1037/xhp0000019

Harter, M. R., Miller, S. L., Price, N. J., LaLonde, M. E., & Keyes, A. L. (1989). Neural processes involved in directing attention. *Journal of Cognitive Neuroscience*, *1*(3), 223–237. https://doi.org/10.1162/jocn.1989.1.3.223

Hodsoll, S., Viding, E., & Lavie, N. (2011). Attentional capture by irrelevant emotional distractor faces. *Emotion*, *11*(2), 346–353. https://doi.org/10.1037/a0022771

Huang, Y. M., Baddeley, A., & Young, A. W. (2008). Attentional Capture by Emotional Stimuli is Modulated by Semantic Processing. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(2), 328–339. https://doi.org/10.1037/0096-1523.34.2.328

Humar, I., Gradisar, M., Turk, T., & Erjavec, J. (2014). The impact of color combinations on the legibility of text presented on LCDs. *Applied Ergonomics*, *45*(6), 1510–1517. https://doi.org/10.1016/j.apergo.2014.04.013

ISO 9241-306:2018. (2018). *Ergonomics of human-system interaction – Part 306: Field assessment methods for electronic visual displays*.

Ko, Y. H. (2017). The effects of luminance contrast, colour combinations, font, and search time on brand icon legibility. *Applied Ergonomics*, *65*, 33–40. https://doi.org/10.1016/j.apergo.2017.05.015

Koster, E. H. W., Crombez, G., Van Damme, S., Verschuere, B., & De Houwer, J. (2004). Does imminent threat capture and hold attention? *Emotion*, *4*(3), 312–317. https://doi.org/10.1037/1528-3542.4.3.312

Lau, N., Veland, Ø., Kwok, J., Jamieson, G. A., Burns, C. M., Braseth, A. O., & Welch, R. (2008). Ecological interface design in the nuclear domain: An application to the secondary subsystems of a boiling water reactor plant simulator. *IEEE Transactions on Nuclear Science*, *55*(6), 3579–3596. https://doi.org/10.1109/TNS.2008.2005979

Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current Directions in Psychological Science*, *19*(3), 143–148. https://doi.org/10.1177/0963721410370295

Lavie, N., Hirst, A., De Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*(3), 339–354. https://doi.org/10.1037/0096-3445.133.3.339

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics*, *56*(2), 183–197. https://doi.org/10.3758/BF03213897

Lee, J., Jung, B., & Chu, W. (2015). Signaling environmental altruism through design: The role of green cue prominence in hybrid cars. *International Journal of Design*, *9*(2), 79–91.

Matthews, M. L. (1987). The influence of colour on CRT reading performance and subjective comfort under operational conditions. *Applied Ergonomics*, *18*(4), 323–328. https://doi.org/10.1016/0003-6870(87)90140-2

McCarley, J. S., & Steelman, K. S. (2013). Visual attention and display design. In *Neuroergonomics: A Cognitive Neuroscience Approach to Human Factors and Ergonomics* (pp. 51–68). Palgrave Macmillan, London. https://doi.org/10.1057/9781137316523_3

McDougald, B. R., & Wogalter, M. S. (2014). Facilitating pictorial comprehension with color highlighting. *Applied Ergonomics*, *45*(5), 1285–1290. https://doi.org/10.1016/j.apergo.2013.05.008

Miran, S. M., Ling, C., James, J. J., Gerard, A., & Rothfusz, L. (2017). User perception and interpretation of tornado probabilistic hazard information: Comparison of four graphical designs. *Applied Ergonomics*, *65*, 277–285. https://doi.org/10.1016/j.apergo.2017.06.016

Muhl-Richardson, A., Godwin, H. J., Garner, M., Hadwin, J. A., Liversedge, S. P., & Donnelly, N. (2018). Individual Differences in Search and Monitoring for Color Targets in Dynamic Visual Displays. *Journal of Experimental Psychology: Applied*. https://doi.org/10.1037/xap0000155

Müller, H. J., Geyer, T., Zehetleitner, M., & Krummenacher, J. (2009). Attentional Capture by Salient Color Singleton Distractors Is Modulated by Top-Down Dimensional Set. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), 1–16. https://doi.org/10.1037/0096-1523.35.1.1

Murray, D. C., & Deabler, H. L. (1957). Colors and mood-tones. *Journal of Applied Psychology*, *41*(5), 279–283. https://doi.org/10.1037/h0041425

Na, N., & Suk, H. J. (2014). The emotional characteristics of white for applications of product color design. *International Journal of Design*, *8*(2), 61–70. https://doi.org/https://doi.org/10.1007/978-3-319-04798-0_20

Naglieri, J. A., & Rojahn, J. (2001). Gender differences in planning, attention, simultaneous, and successive ({PASS}) cognitive processes and achievement. *Journal of Educational Psychology*, *93*(2), 430–437. https://doi.org/10.1037//0022-0663.93.2.430

Nordfang, M., Dyrholm, M., & Bundesen, C. (2013). Identifying bottom-up and top-down components of attentional weight by experimental analysis and computational modeling. *Journal of Experimental Psychology: General*, *142*(2), 510–535. https://doi.org/10.1037/a0029631

Ohman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention_Snakes in the grass.pdf. *Journal of Experiemntal Psychology: General*, *130*(3), 466–478. https://doi.org/10.1037/AXJ96-3445.130.3.466

Öhman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: A threat advantage with schematic stimuli. *Journal of Personality and Social Psychology*, *80*(3), 381–396. https://doi.org/10.1037/0022-3514.80.3.381

Page, T., Thorsteinnson, G., & Ha, J. G. (2012). Using Colours to alter Consumer Behaviour and Product Success. *International Journal of Contents*, *8*(1), 69–73. https://doi.org/10.5392/IJoC.2012.8.1.069

Poulton, E. C. (1975). Colours for sizes: A recommended ergonomic colour code. *Applied Ergonomics*, *6*(4), 231–235. https://doi.org/10.1016/0003-6870(75)90116-7

Ratwani, R. M., Trafton, J. G., & Boehm-Davis, D. A. (2008). Thinking Graphically: Connecting Vision and Cognition During Graph Comprehension. *Journal of Experimental Psychology: Applied*, *14*(1), 36–49. https://doi.org/10.1037/1076-898X.14.1.36

Resnick, M. L., & Albert, W. (2016). The Influences of design esthetic, site relevancy and task relevancy on attention to banner advertising. *Interacting with Computers*, *28*(5), 680–694. https://doi.org/10.1093/iwc/iwv042

Schaie, K. W. (1961a). A {Q}-sort study of color-mood association. *Journal of Projective Techniques*, *25*(3), 341–346. https://doi.org/https://doi.org/10.1080/08853126.1961.10381048

Schaie, K. W. (1961b). Scaling the Association between Colors and Mood-Tones. *The American Journal of Psychology*, *74*(2), 266. https://doi.org/10.2307/1419412

Shive, J., & Francis, G. (2013). Choosing colors for map display icons using models of visual search. *Human Factors*, *55*(2), 373–396. https://doi.org/10.1177/0018720812459341

Snowden, R. J. (2002). Visual attention to color: Parvocellular guidance of attentional resources? *Psychological Science*, *13*(2), 180–184. https://doi.org/10.1111/1467-9280.00433

Starke, S. D., & Baber, C. (2018). The effect of four user interface concepts on visual scan pattern similarity and information foraging in a complex decision making task. *Applied Ergonomics*, *70*, 6–17. https://doi.org/10.1016/j.apergo.2018.01.010

Theeuwes, J. (1992). Perceptual selectivy for color and form. *Percept Psychophys*, *51*(6), 599–606.

Theeuwes, J. (1994). Stimulus-Driven Capture and Attentional Set: Selective Search for Color and Visual Abrupt Onsets. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(4), 799–806. https://doi.org/10.1037/0096-1523.20.4.799

Travis, D. S. (1990). Applying visual psychophysics to user interface design. *Behaviour & Information Technology*, *9*(5), 425–438.

Travis, D. S., Bowles, S., Seton, J., & Peppe, R. (1990). Reading from Color Displays: A Psychophysical Model. *Human Factors*, *32*(2), 147–156. https://doi.org/10.1177/001872089003200203

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Tsai, T. W., Chang, T. C., Chuang, M. C., & Wang, D. M. (2008). Exploration in emotion and visual information uncertainty of websites in culture relations. *International Journal of Design*, *2*(2), 55–66.

Van Laar, D., & Deshe, O. (2002). Evaluation of a visual layering methodology for colour coding control room displays. *Applied Ergonomics*, *33*(4), 371–377. https://doi.org/10.1016/S0003-6870(01)00048-5

Van Laar, D., & Deshe, O. (2007). Color Coding of Control Room Displays: The Psychocartography of Visual Layering Effects. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *49*(3), 477–490. https://doi.org/10.1518/001872007X200111

Vazquez, E., Gevers, T., Lucassen, M., van de Weijer, J., & Baldrich, R. (2010). Saliency of color image derivatives: a comparison between computational models and human perception. *Journal of the Optical Society of America A*, *27*(3), 613. https://doi.org/10.1364/JOSAA.27.000613

Ware, C. (2008). *Visual Thinking: For Design*. *Ergonomics* (Vol. 53). Morgan Kaufmann. https://doi.org/10.1080/00140130903458285

Wexner, L. B. (1954). The degree to which colours (hues) are associated with mood tones. *Journal of Applied Psychology*, *38*(6), 432–435. https://doi.org/https://doi.org/10.1037/h0062181

Wolfe, J. M. (2007). Guided search 4.0. *Integrated Models of Cognitive Systems*, (3), 99–120. https://doi.org/10.1007/978-94-011-5698-1_30

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 419–433.

Wolfe, J. M., & Gray, W. (2007). Guided search 4.0. *Integrated Models of Cognitive Systems*, 99–119.

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, *5*(6), 495–501. https://doi.org/10.1038/nrn1411

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*(3), 58. https://doi.org/10.1038/s41562-017-0058

Zacks, J. M., & Tversky, B. (2003). Structuring Information Interfaces for Procedural Learning. *Journal of Experimental Psychology: Applied*, *9*(2), 88–100. https://doi.org/10.1037/1076-898X.9.2.88

Paper 5:

Attention Affordances: Applying Attention Theory in the Design of Complex Visual Interfaces (submitted), by Andersen, E., Goucher-Lambert, K., Cagan, J., Maier, A.. Submitted to the Journal of Experimental Psychology: Applied.

Attention Affordances: Applying Attention Theory in the Design of Complex Visual Interfaces

Emil Andersen*[1], Kosa Goucher-Lambert[2], Jonathan Cagan[2], Anja Maier[1]

***Corresponding Author**

[1]Management Engineering, Technical University of Denmark, Produktionstorvet, Building 424, 2800 Kgs. Lyngby, Denmark

[2]Mechanical Engineering, Carnegie Mellon University, Scaife Hall, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA

Abstract

The design of visual interfaces plays a crucial role in ensuring swift and accurate information search for operators who use procedures and information tables to cope with problems arising during emergencies. The primary cognitive mechanism involved in information search is visual attention. However, the design of interfaces is seldom done through applying predictions of attention theory. Conversely, theories of attention are seldom tested in applied contexts. Combining application and attention theory thus stands to benefit both areas. Therefore, this study tested three theories of visual attention that are especially relevant for information processing in emergencies (Feature Integration Theory, Load Theory and Dilution Theory) and predictions about attentional guidance and capture of color in a complex visual interface. Evidence was found for several predictions from theory, especially from the Feature Integration Theory. Implications for design practice and attention theory are discussed.

Keywords: Visual Attention, Affordances, Cognitive Load, Visual Load, Color, Human-centered design, Human Performance

In complex socio-technical environments, the performance of human operators heavily depends on the design of the visual aids at their disposal. For example, nuclear power plant operators depend on the critical information presented in several complex interfaces to inform their decisions. If the correct information is found slowly due to suboptimal design, decision-making may suffer and cause severe adverse effects. Similar practice areas include aviation, military strategy and medicine, all sharing in common that the correct information must be found in a complex visual interface at high speed.

Visual interfaces must therefore be designed such that they allow for the correct actions to be performed. The possible actions that an interface allows, commonly referred to as affordances (Gibson, 1978; Norman, 1988), must thus be managed by the designer to increase the likelihood of the correct actions being performed. A useful tool for managing affordances lies in user studies, which improve design through increased knowledge of the potential user (Crilly, Moultrie, & Clarkson, 2009). To this end, previous research has provided guidelines for practitioners on how to improve visual performance with regards to increasing, e.g., general aesthetic appeal (Blijlevens, Mugge, Ye, & Schoormans, 2013; Choi, Orsborn, & Boatwright, 2016; Crilly, Moultrie, & Clarkson, 2004; Orsborn, Cagan, & Boatwright, 2009) or through investigation of specific product tests and user evaluation studies (Braseth & Øritsland, 2013; Karlsson, 2007; Lau et al., 2008; Na & Suk, 2014; Ranscombe, Hicks, Mullineux, & Singh, 2012; Weyer et al., 2010). However, such studies seldom relate their findings to the underlying cognitive mechanisms, nor aim to towards generalisable theories of interactions with specific design features. Conversely, the present study aims at making guidelines for improving visual design for emergencies by using theories of attention, which is the primary cognitive mechanism responsible for determining what information reaches consciousness (Baars & Franklin, 2007; Dehaene, Charles, King, & Marti, 2014).

A central feature of the attention system is that its capacity is limited, allowing only three to five objects to enter consciousness at any given time (Deutsch & Deutsch, 1963; Sperling, 1960). As a result, its allocation is especially important for complex interfaces that contain large amounts of information. Through highly focused and controlled experimental studies, researchers in fields such as cognitive psychology and neuroscience have developed topic specific theories (e.g. in relation to visual load, Lavie & Tsal, 1994), as well as generalized theories of attention allocation (e.g. Guided Search, Wolfe et al., 1989; The Theory of Visual Attention, Bundesen, 1990) that have proven to robustly predict the allocation of attention in highly controlled contexts. However, these findings have seldom been tested in relation to concrete use-cases. A potential reason for why they are seldom translated is that the amounts of influencing variables in real world scenarios usually vastly outnumber the influencing variables investigated in the highly controlled experiments: When shaping visual aids for use in critical situations, a designer must consider *all* aspects of attention simultaneously. However, due to the lack of studies that combine both applied scenarios and information about the underlying cognitive mechanisms, little information is available regarding how the various aspects of attention interact with each other in more realistic visual design cases. This paper attempts to address this gap in information by investigating how participants perform when searching for information in a display that simultaneously mimics a complex visual interface used in nuclear control rooms and the interfaces used in experimental psychology to allow for predictions about attention allocation. Thereby, insights are gained on how predictions and insights from attention research can be applied when using design to communicate visual information in emergency scenarios.

This paper first reviews research and frameworks from ecological psychology and design research that are useful for understanding how interfaces are used. It is argued that designers create interfaces with certain actions in mind by creating products that afford those types of actions, and that user studies may lead to better understand which design options afford specific actions. Second, research on visual attention is outlined that provides insights into which visual components afford certain types of attention allocation. The focus is on theories that have relevance for emergencies. Third, two visual search experiments are conducted that test multiple predictions from the attention literature in a complex visual interface based on those used in nuclear power plant control rooms. Fourth and finally, the results and their implications for design practice and attention theory are discussed. The following sections show the relation between user studies and improved design, as illustrated by the alignment of intended and actual affordances. In relation to this, Load Theory (Lavie, Hirst, De Fockert, & Viding, 2004; Lavie & Tsal, 1994), Feature Integration Theory (Treisman & Gelade, 1980) and Dilution Theory (Benoni & Tsal, 2010; Tsal & Benoni, 2010) are outlined due to their possible applicability to interface design.

Aligning Designer Intention and Actual Use

One of the main goals of this work is to offer suggestions that can improve the design of visual interfaces for use in emergencies, based on attention theory. To this end, it is important to first consider how information from user studies can aid designers in developing better solutions. Previous research has shown that designers hold distinct intentions with regards to how the product being created will be used. Research has also demonstrated that designers exert effort towards making the design in a way that increases the likelihood of the intended use also being reflected in the actual use (Crilly et al., 2009). This reflects design as a communicative process that involves a designer telling the user *how* to use the product *through* the product (Crilly, Good, Matravers, & Clarkson, 2008; Crilly, Maier, & Clarkson, 2008). This paper takes this design-as-communication approach as its basic assumption with regards to how to improve design. Under this assumption, an important way that designers can increase the likelihood of a device's intended use becoming its actual use is through manipulating the possible actions of the designed object. These possible actions are commonly referred to as *affordances*, usually referencing to the original formulation by Gibson (1978) and/or the application to design by Norman (Norman, 1988). Adopting this concept, a product is considered to have a number of affordances, which denotes the possible actions that can be performed in relation to or with the product. Furthermore, the concept usually relates to likelihood of use (e.g., a chair more readily affords sitting on it than throwing it), whereby a metric for success of a designed object lies in the extent to which the action that was intended for the product is also the action that is most readily afforded by the item. Previous research has shown that designers can increase the likelihood of their intended affordances matching actual affordances by performing user studies (Crilly et al., 2009; Maier, Fadel, & Battisto, 2009): A prototypical user study will thus elucidate the likelihood of specific affordances of a product pertaining to specific groups. Applying such user or group knowledge is sometimes referred to as user-entered design or human-centered design (e.g. Cagan & Vogel, 2001, for a recent review see Boy, 2017). Such user-centered design studies have resulted in insights relating to topics such as aesthetic preferences and user experience (Blijlevens et al., 2013; Hung & Chen, 2012; Ranscombe et al., 2012; Tovares, Boatwright, & Cagan, 2014), perceived sustainability and its relation to preference (Goucher-Lambert & Cagan, 2014; Reid, Gonzalez, & Papalambros, 2010), and the emotion evoked by products and their specific attributes (Boatwright & Cagan, 2010; Demir, Desmet, & Hekkert, 2009; Desmet, 2012; Goucher-Lambert, Moss, & Cagan, 2017; Karlsson, 2007). Alternatively, more recent evidence has shown that designs can be improved by relying on findings that are common to all humans due to our shared phylogenetic heritage, such as basic principles object coherence (Leder, Carbon, & Kreuzbauer, 2007; Nørager, 2009) or gestalt laws (Lugo, Schmiedeler, Batill, & Carlson, 2016), or very large cultural norms, such as 'Westerner' interpretation of colors (Aslam, 2006). These studies represent an increasing body of knowledge on the affordances of specific properties. While the original definition of affordances is unconstrained, previous research has indeed mostly considered the concept with regards to concrete actions (e.g., the above chair example), aesthetics (Xenakis & Arnellos, 2013), semantics (You & Chen, 2007), tool usage (Wagman & Carello, 2003), interface design (Stefanucci, Creem-Regehr, Thompson, Lessard, & Geuss, 2015) or to fields such as engineering design (Ciavola & Gershenson, 2016) or design and architecture (Maier et al., 2009). Recently, however, Still & Dark (2013) proposed that affordances can represent any automatic cognitive process that is evoked when viewing an object, and that the relationships can be both evolutionary and culturally acquired. They conclude by speculating that in the presence of several perceived affordances, designers may benefit from models of basic cognitive mechanisms (e.g., the biased competition model of Desimone & Duncan, 1995) to inform their designs.

This paper takes this notion a step further, showing how affordances of a visual interface can be related to, and predicted through, the application of attention theory. Understanding the visual attention system is particularly important for predicting what information will receive attentional processing, or, in other words, what information visual interfaces most readily afford being processed. Drawing upon these insights thereby represents an important avenue for increasing the likelihood that the intended affordance of a display is also the most likely (Kozine, 2007; McCarley & Steelman, 2013). In the following section, the foundational studies and theories that form the basis of most of attention research are outlined. Following this, theories that relate to, and give concrete predictions about, attention allocation are described and the predictions are tested using a complex visual interface that mimics those typically found in nuclear control rooms.

The Guidance and Capture of Attention

A central topic of research on the visual attention system is to what extent attention is under our control. Several theories of attention (Bundesen, 1990; Wolfe, 2007; Wolfe et al., 1989; Wolfe & Horowitz, 2017) have proposed that which objects catch our eyes depends on both their relevance for our goals (referred to as 'top-down' weight) and the inherent relevance of the objects' properties (also referred to as 'bottom-up' weight). For example, top-down control of attention results means that if you are scanning for a word in capital letters in a search display, your attention will be more likely to be allocated to capital words in general, while bottom-up control of attention means that a blinking word may catch your attention regardless of your goal due to the inherent capturing properties of newly appearing stimuli (Theeuwes, 2010; Wolfe & Horowitz, 2017). A prevalent illustration of the interaction between the effects of bottom-up and top-down features is that the objects 'race' towards attention with various speeds based on their top-down and bottom-up relevance (Bundesen, 1990; Desimone & Duncan, 1995). Additionally, several conditions have been found to modify the degree to which the attention system is guided and captured by top-down and bottom-up features. In the following sections, the effects of load, dilution and color are outlined due to their relevance to design of interfaces used in critical situations. For each collection of findings, the specific predictions are summarized, which serve as the basis for the two visual search experiments presented in this paper.

Load Theory and Feature Integration Theory: Cognitive and Visual Load

In both everyday and emergency settings, users of visual interfaces are subject to varying degrees of cognitive strain. Whether it is because of having to perform several tasks at once, having to hold varying amounts of information in memory, or varying amounts of multimodal stimulation (such as an alarm going off), the cognitive load of any user will be under varying demands depending on the situation. The Load Theory of Selective Attention and Cognitive Control (henceforth Load Theory, Lavie, 2005, 2006, 2010; Lavie et al., 2004; Lavie & Tsal, 1994) predicts that when cognitive load goes up, the ability to ignore distracting stimuli goes down, due to a decrease in available cognitive inhibitory resources. Given the natural variations in workload in users of visual interfaces in the real world, it seems highly relevant to see if the predictions of Load Theory extend to a more realistic interface.

In addition to the cognitive load induced from the surroundings, the interface itself can induce varying levels of load. In attention theory, load induced by the amount and characteristics of the objects in the visual scene is most commonly referred to as visual load, again usually with reference to Load Theory (Lavie et al., 2004; Lavie & Tsal, 1994). The most commonly referred examples of high and low visual load stem from the experiments that form the basis of the Feature Integration Theory (Treisman & Gelade, 1980). In a classic low visual load display, the target is separable from the distractors by a single feature. For example, it can be the only X amongst a number of O's, or the only blue object amongst a number of red objects. In such situations, a high-capacity pre-attentive processing stage can aid the visual search, leading to a fast, parallel search that is only marginally affected by the number of distracting objects in the display. In a classic high load display, the target and distractors are all unique members of the same category. For example, the target can be an X with distracting stimuli L S H, or a blue target amongst red, green and orange distractors. In such situations, the pre-attentive processing is not sufficient for attention allocation, and instead the observer must inspect each element and determine whether it was the target in a slow, serial fashion. In sum, high visual load requires slow, individual processing of each object, where low visual load allows for rapid processing of several objects in parallel. The difference in visual load has thus been shown to be highly influential for visual search speed in a highly controlled setting. Furthermore, Treisman & Gelade's theory shown promise for predicting visual search for signs, symbols an icons (McDougall, Tyrer, & Folkard, 2006) and for targets in color-coded and intensity-coded displays (Yamani & McCarley, 2010). Additionally, the effect of visual load has been shown to extend to how distractors are processed. According to Load Theory, higher visual load will reduce the processing of distractors, as all capacity will be occupied by target-relevant objects. Conversely, low visual load affords spare processing capacity to spill over to irrelevant objects, leading to higher processing of distractors (Lavie et al., 2004; Lavie & Tsal, 1994). This finding has been shown to extend to irrelevant, real-life objects cartoon characters (Forster & Lavie, 2008), although these effects have recently been shown to be limited in their generalizability (Lleras, Chu, & Buetti, 2017). Taken together, these theories thus make specific predictions with regards to how users of complex displays should be able to filter relevant visual objects from irrelevant visual objects, as well as the likelihood that they are distracted.

Visual Dilution Theory

As a competing theory to Load Theory, Visual Dilution theory (Benoni & Tsal, 2010; Tsal & Benoni, 2010; Wilson, Muroi, & MacLeod, 2011) addresses distractor processing under visual load as well, but alters the theory to take into account the total amount of objects in the visual display. Drawing on the findings of Stroop Dilution (Kahneman & Chajczyk, 1983), dilution theory of distractor interference states that processing of distractors depends on the total amount of objects in the display: If there are more objects, then distractors will have lower influence because their effect is diluted by even entirely irrelevant objects. This effect has been explained in some different ways. Kyllingsbæk et al. (2011) ascribe the effect to a lowered probability of attention being allocated to any given object. While distractors may still have a larger probability than an irrelevant distractor, the dilution effect results in overall lower probability of it being selected. Based on this, while Dilution Theory does not address whether irrelevant objects will influence processing speed overall, the theory predicts that distractors should have a lower effect with more irrelevant objects. The predictions of this theory thus could be highly influential for understanding how easily a user is distracted by irrelevant objects in a complex display.

Color Guidance

Color is a common and useful tool for distinguishing between objects in complex visual displays (Jameson, Kaiwi, & Bamber, 2001; Spence & Efendov, 2001; Spence, Kutlesa, & Rose, 1999; Vazquez, Gevers, Lucassen, van de Weijer, & Baldrich, 2010; Ware, 2008). This is in concordance with attention theory, which has repeatedly shown that color can both guide (Wolfe, 2007; Wolfe & Horowitz, 2017) and capture (Nordfang, Dyrholm, & Bundesen, 2013; Theeuwes, 1992, 1994) attention. While these studies mostly dealt with color as a category, a recent study extended this to show that individual colors have individual effects on attention Andersen & Maier (2017). Furthermore, previous studies have shown that the overall and individual colors depend on visual load (Biggs, Kreager, & Davoli, 2015; Andersen & Maier, under review). Additionally, previous research has also shown that search in visual displays using colors to distinguish between objects depend heavily on the combination of the individual colors (Francis, Bias, & Shive, 2010; Müller, Geyer, Zehetleitner, & Krummenacher, 2009; Shive & Francis, 2013; Starke & Baber, 2018). Therefore, applying attention theory on the guidance and capture of color is important when judging the attention affordances of complex displays.

Overview of Experiments

Two experiments were created to test the outlined predictions of the Feature Integration Theory, Load Theory, and Dilution Theory, as well as the presented findings on attentional guidance by color, within a design relevant context. The visual interfaces used were created to mimic a simplified nuclear control room interface (based on the illustrations of Braseth & Øritsland, 2013), while simultaneously sharing enough characteristics with the displays used in the attention literature to allow testing the above predictions. The two experiments were identical in all aspects except that in Experiment 1, participants were given information about both the target color of an object (a barometer) and its target number, whereas the target color was unknown in Experiment 2.

Experiment 1 thus mimics the common methodology for studying attentional guidance of color (e.g. Biggs et al., 2015; Wolfe, Jeremy M.; Cave, Kyle R.; Franzel, 1989; Wolfe & Horowitz, 2004, 2017), and Experiment 2 mimics the common methodology for studying attentional capture of color (e.g. Nordfang et al., 2013; Theeuwes, 1992). In concurrence with Nordfang et al., (2013), faster search times were expected for colored targets in both experiments, but more so when participants are given information about the target color. The remaining individual hypotheses associated with each prediction and the manipulations employed to test them are outlined in the sections below and summarized in Figure 1.

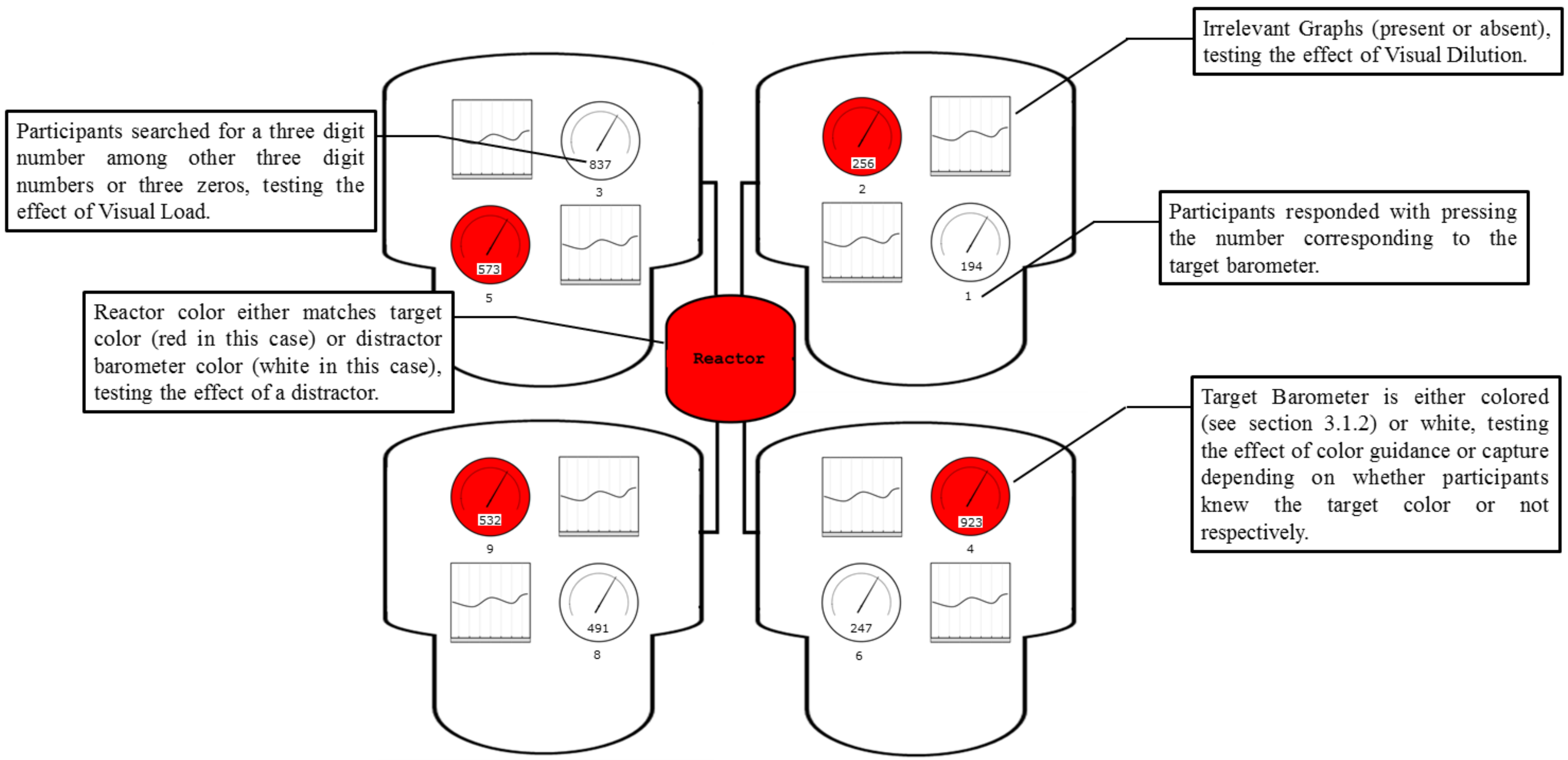


Figure 1: An example target search display with explanations for the various manipulations and their relation to theory. See the online article for a colored version of this figure.

Experiment 1: Guidance of Color

Method

Participants. Twenty-five (nineteen female) students at Carnegie Mellon University and University of Pittsburgh participated in the experiment in exchange for monetary compensation. Participants were between 18 and 30 years old, and were required to have normal or corrected-to-normal vision, not be color blind, not suffer from ADHD or depression, or have a family history of ADHD or depression. Participants were pre-screened for color blindness prior to the experiment.

Overall Design. Visual interfaces were created that mimic a simplified nuclear control room interface based on the illustrations of Braseth & Øritsland (2013), while simultaneously keeping relevant objects equidistant from the center of the display. All interfaces contained a “Reactor” object (of varying color, see Figure 2), which measured 2.6˚diagonally and 3.5˚ horizontally, connected to four “Steam Generators”, which measured 8.7˚ horizontally at the widest point, 6.0˚ horizontally at the narrowest point and 6.9˚ diagonally. Each “Steam Generator” contained two contained two “Barometers”. The “Barometers” varied in color (see Figure 2 for the various colors), measured 1.9˚ in diameter, and contained a white box with a three digit identification number, which participants used for visual search. Below each barometer was a single digit for participants to report in the search task. This mimics the design of Theeuwes (1992) and Nordfang et al., (2013) in that that the target color was not part of the target number. In half of the interfaces, each “Steam Generator” also contained two “Trend Graphs”, measuring 1.9˚ both horizontally and diagonally, to increase the visual dilution. “Barometers”, “Reactors”, “Steam Generators” and “Trend Graphs” were created in Microsoft PowerPoint 2010. Text objects were created in E-Prime and used the Courier New font. Colored objects used the “standard colors” Red, Light Blue, Light Green, Yellow, Orange and Purple from Microsoft PowerPoint 2010 in concordance with (Andersen & Maier, 2017; Andersen & Maier, under review). All text objects used the Courier New font, and were created in E-Studio version 2 for Windows. Figure 1 shows an example interface with the various manipulations and their relation to theory.

Colour	R	G	B	Contrast	Luminance
Red	255	0	0	4	29
Yellow	255	255	0	1.1	106
Green	0	176	80	2.9	33
Blue	0	112	192	5.1	15
Orange	255	192	0	1.6	66
Purple	112	48	160	8	7

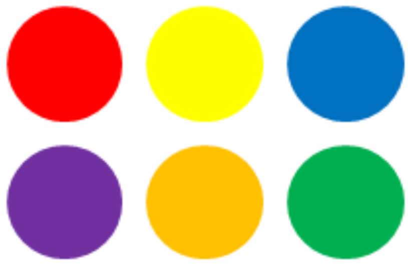


Figure 2: Color codes, contrast scores, luminance scores, and visual representation of the used colors. See the online article for a colored version of this figure.

Cognitive Load. Cognitive load was manipulated through a modified version of the dual task paradigm of Lavie et al., (2004), which uses a reduced version of Sternberg's (1966) short-term recognition task to manipulate cognitive load while participants perform a search task. This experiment varied in that participants remembered letters and searched for numbers (whereas the converse was used by Lavie et al., 2004). Participants were asked to remember a letter sequence of one (low load condition) or five (high load condition) letters. Letters were consonants (y was considered a vowel) to avoid reduced load from chunking. The sequence was presented for 500ms in the low load condition and 2000ms in the high load condition similar as in Lavie et al., (2004). After this, the participants performed the search task, and were then presented with a single letter probe, which they determined to be part of the sequence or not. There was a 50-50 chance that the presented probe was present or absent. Participants pressed the CTRL key if the target was present and the ALT key if the target was absent. For each trial, participants were given feedback on their performance on the memory task in that specific trial, as well as all trials so far (excluding practice). In concordance with Load Theory, higher overall reaction times were expected under higher load, and larger distractor interference.

Visual Load. Visual load was manipulated through conditions that allowed participants to search in parallel or serially for the number target, to see the effect on overall reaction time and distractor processing. In the high load condition, all eight barometers have a sequence of three random numbers (e.g., 957). In the low load condition, the target is the same as in the high load condition and, but all other barometers have three zeros (000) as their number sequence. The experiment thus mimics the common methodology for studying the effect of visual load (e.g. Biggs, Kreager, & Davoli, 2015b; Lavie et al., 2004; Lavie & Tsal, 1994). Higher visual load was expected to lead to higher reaction times.

The effect of distractor interference under load was tested by varying the color of the central “Reactor Object” to be either congruent (i.e., the same as the target) or incongruent (i.e., white, the color of the non-target barometers) with the target color. Participants were told to ignore the distractor in all cases. This experiment thus mimics the flanker response-competition task (Eriksen & Eriksen, 1974), with a central flanker (Beck & Lavie, 2005; Wilson et al., 2011). Differences in response times between the congruent and incongruent conditions were considered to indicate distractor interference. In concordance with Load Theory, lower distractor interference was expected in the high load condition. Furthermore, in concordance with Biggs et al., (2015) and Andersen & Maier (under review) larger effects of guidance and capture of color under higher visual load were anticipated.

Visual Dilution. Visual dilution was manipulated through the inclusion or exclusion of irrelevant “Trend Graphs” in the “Steam Generators”. Participants were instructed to ignore them whenever they appeared. In concordance with the Dilution (Benoni & Tsal, 2010; Tsal & Benoni, 2010; Wilson et al., 2011) account and Theory of Visual Attention (TVA, Kyllingsbæk et al., 2011), lower distractor interference was expected under higher visual dilution.

Placement. The possible effect of reading order was controlled for by randomizing the placement of both target and distractor objects across participants. In concordance with previous findings (Buscher, Cutrell, & Morris, 2009; Buscher, Dumais, & Cutrell, 2010; Cutrell & Guan, 2007) participants were expected to find targets faster in the top-left quadrant of the display due overtraining for the left-to-right reading order.

Apparatus. The experiment was conducted on a laptop computer with an externally connected screen and keyboard. Stimuli were displayed using E-Studio version 2 for Windows.

Experiment Procedure. Each participant completed one training block (20 trials) and three experiment blocks (102 trials each for 306 trials total). Each trial proceeded as follows: First, a screen appeared for 2000ms or 500ms (see section 3.1.4 for details), which showed the letter sequence to be remembered, followed by a 30ms visual mask. Second, the target barometer, including color and barometer number, was shown for 2000ms followed by a 30ms visual mask. Third, a fixation cross appeared in the middle of the screen for 500ms to ensure that participants were focused in the middle of the display at the start of each search task. Fourth, the search display (see Figure 1 for an overview of the variations) appeared until the participant responded with the corresponding number using the numpad (maximum response time allowed was 8000ms based on pilot tests) had passed. Fifth, participants were shown a letter probe, and pressed CTRL if the number was present in the memory sequence or ALT if the number was absent (see section 3.1.4 for details) (maximum response time allowed was 6000ms based on pilot tests). Sixth and final, participants were given feedback on their performance in the memory task for 1000ms. Figure 3 shows the experiment procedure.

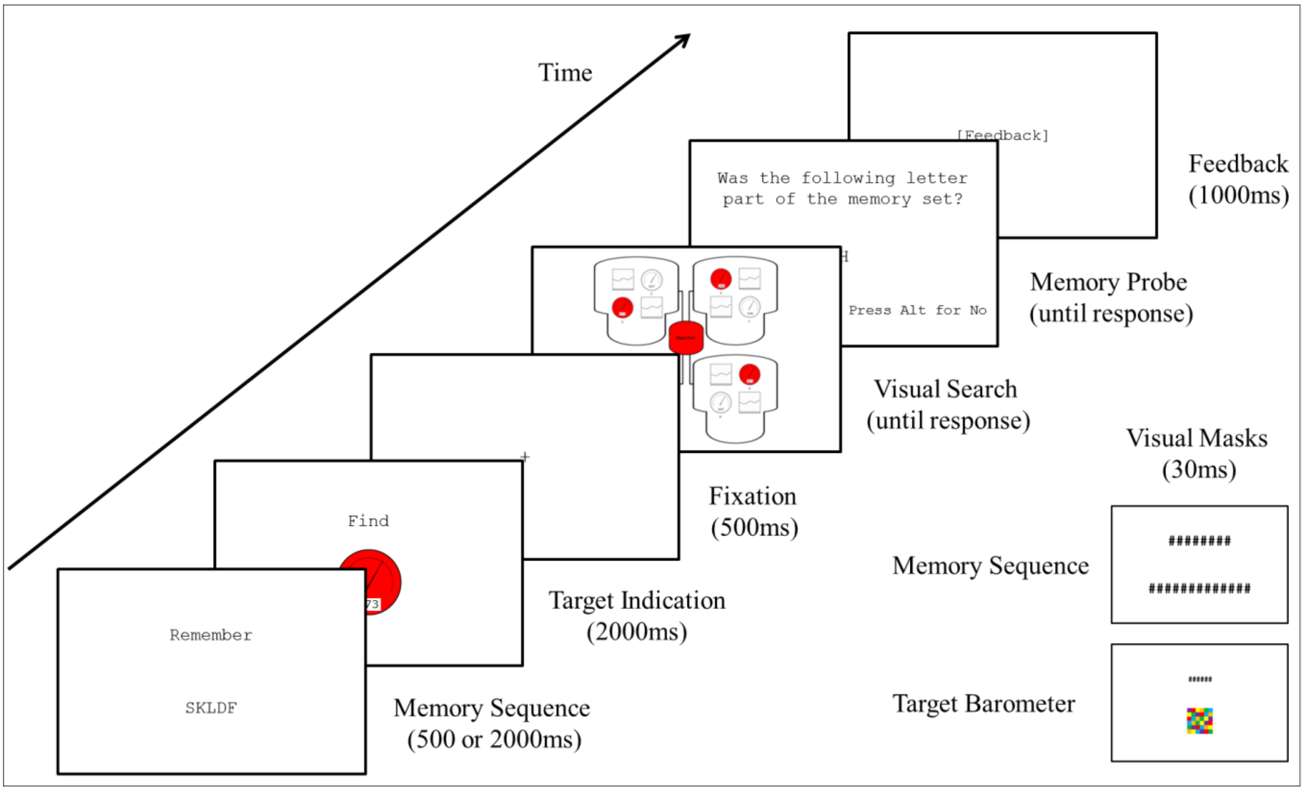


Figure 3: The experiment procedure. See the online article for a colored version of this figure.

Results

Analysis was conducted in R studio for Windows using base R and the nlme package. The dependent variables were reaction time (RT) and accuracy for the visual search task and memory task. The independent variables were target color, distractor color, distractor interference (measured as the difference between distractor types), cognitive load, visual load, visual dilution and placement. A linear mixed effects model was fit to the data, and results were evaluated using a Repeated-Measures ANOVA model with Subject modelled as a random factor. Reaction time data were log-transformed prior to analysis to better fit the assumption of normal distribution. An overview of the results is given in Table 1

Table1: Mixed Effects Linear Model for Experiment 1

Variable	F-value	p-value
Target Color	72.86	<.0001**
Distractor Color	1.78	0.0994
Distractor Type	5.15	0.0233*
Cognitive Load	32.13	<.0001**
Visual Load	98.63	<.0001**
Irrelevant Distractor Count	0.17	0.6819
Target Position	267.34	<.0001**
Target Color x Visual Load	1.69	0.1193
Distractor Type x Visual Load	0.32	0.5731
Distractor Type x Cognitive Load	0.32	0.5695
Distractor Type x Irrelevant Distractor Count	6.13	0.0133*

*significant **highly significant

Target color was significant ($F = 72.85$, $p<0.0001$) indicating a significant difference in reaction time as a function of individual colors. Post-hoc tests revealed that this difference was due to white targets being found significantly slower than targets of other colors.

There was no significant difference between individual distractor colors ($p > 0.05$). However, there was significant distractor interference ($F= 9.08$, $p=0.0026$), as targets were found significantly slower when the target color was the same as the distractor color. Higher cognitive load significantly lowered search times ($F = 32.10$, $p<0.0001$). Higher visual load significantly increased search times ($F = 98.77$, $p<0.0001$). There was no significant effect of visual dilution ($p > 0.05$), meaning that the presence of the irrelevant graphs did not affect overall reaction time. However, there was a significant interaction between the presence of irrelevant distractors and distractor interference (F = 6.13, p = 0.013), indicating that distractor interference was higher when the irrelevant graphs were absent. Figure 4 shows this relationship. There was a significant effect of placement (F = 267.34, p < 0.0001), with targets being found faster if they were in the top-left quadrant (see Figure 5). Finally, in opposition to what was predicted by Load Theory, there were no significant interactions between cognitive load and distractor interference or visual load and distractor interference. Furthermore, there was no significant interaction between target color and visual load as opposed to previous findings (Andersen & Maier, under review).

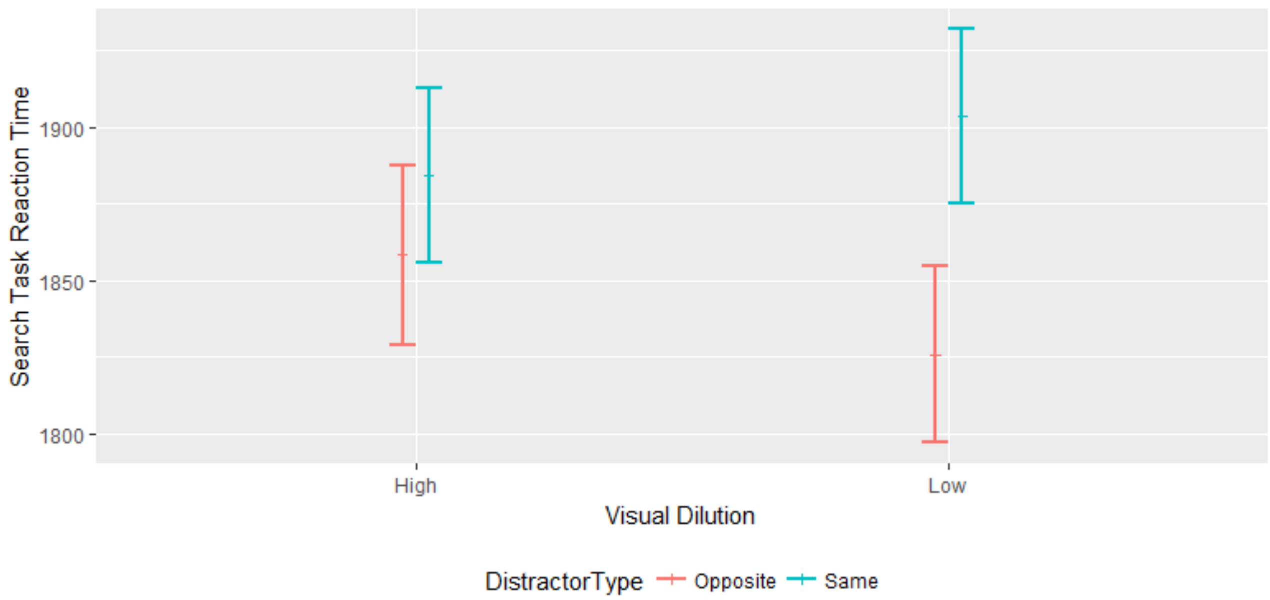


Figure 4: The interaction between Visual Dilution and Distractor Interference. When the diluting graphs are present, the color of the distracting reactor object has no significant effect. When the diluting graphs are absent, the color of the distracting reactor object significantly slows search if it matches the target color. Error bars indicate 95% Confidence Intervals. See the online article for a colored version of this figure.

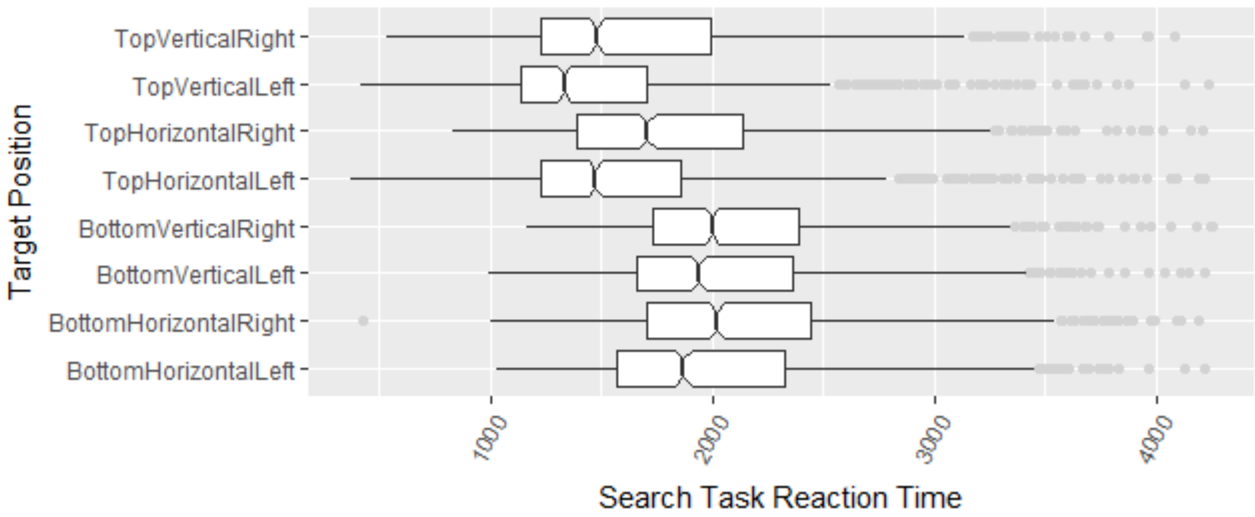


Figure 5: Boxplots for the effect of placement. Items in the two top-left positions are found significantly faster. Notches in the boxplots indicate 95% Confidence Intervals.

Discussion

Experiment 1 studied visual search performance for three-digit numbers that were located in barometers that participants knew the color of. The results showed both congruence and incongruence with the predictions from attention literature. Most notably, the results supported the predictions of Feature Integration Theory (Treisman & Gelade, 1980), namely that participants were faster when they were able to use a single feature to filter out irrelevant objects. This was reflected in a significant effect of visual load, indicating that when participants could use the shape of the number, rather than needing the specific number identity, to conclude their search. Furthermore, there was no significant effect of visual dilution, indicating that participants were unaffected by the presence of the distracting graphs because they were able to filter them out. This, perhaps surprising, result indicates that a cluttered visual display will not be harder to navigate if targets are sufficiently distinguishable by category.

Furthermore, these results clearly show that colored targets were found faster than white targets, in congruence with previous research (Nordfang et al., 2013; Theeuwes, 1992, 1994; Wolfe, 2007; Wolfe & Horowitz, 2017). This indicates that non-white targets had a higher priority in the attention system, which may be due to the higher contrast between the background and the targets (Irwin, Colcombe, Kramer, & Hahn, 2000), given that the results did not show differences between the individual non-white colors. Furthermore there was no significant interaction between individual colors and load, which stands in opposition to previous findings (Biggs et al., 2015; Andersen & Maier, under review), indicating that the other effects took priority for attention allocation in this experiment.

Adding to this, there was a clear reading-order effect, showing that participants found targets in the top-left quadrant faster. The reading-order effect is usually corrected for in attention experiments (as was the case in this experiment), but the results corroborate previous findings (Buscher et al., 2009, 2010; Cutrell & Guan, 2007) in showing its significance for visual interface design. Namely, that high priority targets should be located in the top-left quadrant.

Moving to the second-order effects, the results showed that the effect of the salient distractor object (the reactor in the middle of the display) was consistent with the predictions of Dilution Theory, and inconsistent with Load Theory. Experiment 1 thus corroborates the recent criticism of Load Theory, suggesting that the interaction between Load and distractor processing should instead be attributed to dilution (Benoni & Tsal, 2010). Namely, the results showed that the distracting reactor object only had a significant effect on performance when the irrelevant trend graphs were not present. While the irrelevant graphs did not affect overall performance, the higher visual clutter thus nevertheless reduced the effect of a salient distractor, further supporting the counter-intuitive finding that more complex displays may not be harder to navigate if the target characteristics are well known.

Finally, the results showed a counter-intuitive effect of cognitive load, as higher cognitive load was associated with faster search speed. Post-hoc test revealed that this may have been due to a small (~1.5%), but statistically significant ($F = 14.9$, $p<0.0001$), difference in accuracy. As such, when working memory was beyond the normal maximum capacity, participants were able to find the targets faster at a slight cost of accuracy. A non-significant interaction effect between cognitive load and the distractor type indicates that the faster search speed was not due to lower distractor processing. As this effect is counter to the predicted, further research is needed to clarify this effect. While there were no significant differences between the individual target colors, particularly between white and non-white colors, the results may have been influenced by participants knowing the target color. In real life, however, users of visual interfaces may not always know the target color in advance, but rather only an identifier such as an identification number or shape. To test whether results were different when participants did not know the target color, a second experiment was conducted where participants did not know the target barometer color.

Experiment 2: Capture of Color

Method

Participants. A new sample of twenty-five (sixteen female) students at Carnegie Mellon University and University of Pittsburgh participated in the experiment. Inclusion criteria and compensation were identical to Experiment 1.

Overall Design & Procedure. Experiment 2 was identical to Experiment 1 in all aspects except that participants were only instructed in the target number, as opposed to both the target color and target number in Experiment 1. Given that participants had no knowledge of the target color, distractor interference was measured only in relation to differences between the individual distractor colors.

Results

Due to participants not knowing the target color, the effect of the distracting reactor object's color matching the target color was in itself not a meaningful measure. Instead, it was measured whether there was a significant difference between the individual colors, irrespective of whether they matched the target color or not. With this exception, the analysis was identical to the one conducted in Experiment 1. An overview of the results are shown in Table 2

Table 2: Mixed Effects Linear Model for Experiment 2

	F-value	p-value
Target Color	19.94	<.0001**
Distractor Color	4.09	0.0004**
Distractor Type	3.23	0.0723
Cognitive Load	10.67	0.0011**
Visual Load	135.50	<.0001**
Irrelevant Distractor Count	1.17	0.2797
Target Position	262.73	<.0001**
Target Color x Visual Load	2.52	0.0194*
Distractor Color x Visual Load	0.15	0.9887
Distractor Color x Cognitive Load	0.35	0.9102
Distractor Color x Irrelevant Distractor Count	1.49	0.1763

*significant **highly significant

There was a significant effect of target color ($F = 19.94$, $p<0.0001$) indicating a significant difference in reaction time between individual target colors. Post-hoc tests revealed that this difference was due to white targets being found significantly slower than all colored targets except red, whereas red and purple were found significantly slower than yellow ($p<0.001$), orange ($p<0.01$) and blue targets ($p<0.001$).

There was a significant effect of distractor color ($F = 3.23$, $p< 0.001$) indicating that there was a significant difference between individual distractor colors. Post-hoc test revealed that this was due to red distractors capturing significantly more attention than blue ($p=0.02$), orange ($p=0.034$) and white (p<0.0001) distractors, and purple distractors capturing significantly lower attention than white distractors (p < 0.001). Higher cognitive load significantly lowered search times (F = 10.71, p<0.0011). Higher visual load significantly increased search times (F=135.85, p<0.0001). Furthermore, there was a significant interaction between visual load and target color (F = 2.52, p = 0.0196), indicating a larger difference between colored targets at high visual load, and a larger difference between colored and white targets at low visual load. Figure 6 shows this relationship.

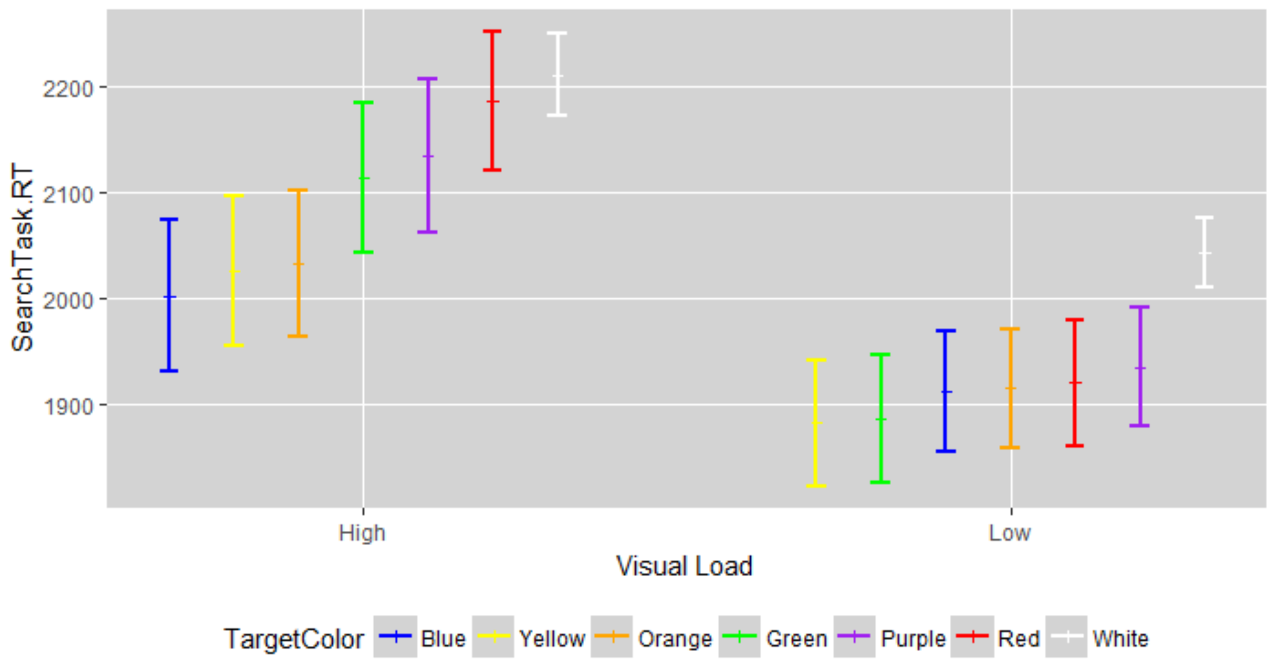


Figure 6: The interaction between Visual Load and the difference between individual target colors. At high load, the individual difference between colors is larger, but the difference between red and white is smaller. At low load, the difference between non-white colors is smaller, but the difference between non-white and white targets is larger. Error bars indicate 95% Confidence Intervals. See the online article for a colored version of this figure.

There was no significant effect of visual dilution (F=1.17, p = 0.28), meaning that the presence of the irrelevant graphs did not affect overall reaction time. Finally, there was a significant effect of placement (F = 263.02, p < 0.0001), with targets being found faster if they were in the top-left quadrant (see Figure 7).

Figure 7: The effect of placement. Items in the two top-left positions are found significantly faster. Notches indicate 95% Confidence Intervals.

Discussion

Experiment 2 studied visual search performance when participants did not know the target color in a complex display. As with Experiment 1, the results showed both congruence and incongruence with the predictions from attention literature, but notably, the congruencies and incongruencies differed from those in Experiment 1.

Experiment 2 reproduced the findings from Experiment 1: that targets in the top-left quadrant were found faster, that irrelevant graphs did not affect search, that visual load slowed search, and that cognitive load improved search speed. As in experiment, post-hoc analysis showed that the difference in speed as a function of cognitive load may have been due to a small (~1.5%), but statistically significant (F = 39.72, p < 0.0001), difference in accuracy. These findings thus further corroborate the prediction by Treisman & Gelade (1980) that participants were able to ignore irrelevant stimuli by only scanning for barometer-like objects, and non-000 target numbers when applicable. Furthermore, the findings indicate that the culturally induced reading order is highly relevant for attention allocation, regardless of the extent of knowledge about the target. Finally, the results of Experiment 2 further underline a need for conducting additional research to elucidate the effect of cognitive load on performance.

However, unlike Experiment 1, the results of Experiment 2 showed a significant difference between individual target and distractor colors. The difference between individual colors was furthermore larger at high visual load, in concordance with previous research (Biggs et al., 2015; Andersen & Maier, under review). Furthermore, the difference between distracting colors was not affected by load, in concurrence with a former study with Andersen & Maier (under review). These results thus indicate that individual colors of both target and distractor objects have individual effects on attention when the target color is unknown. Therefore, when users are expected to be naïve to the expected color scheme (e.g., due to a lack of training opportunities or infrequent exposure to the display) the exact colors used have a significant effect on attention allocation, whereas this was not the case for interfaces where the target color was known.

General Discussion

Two experiments were conducted in order to test predictions from attention theory about how complex visual displays afford the allocation of attention. This was done using variations of a display that mimicked both the interfaces used in nuclear control rooms and those used in experimental psychology. Specifically, it was investigated whether attention theory could predict the attention affordances of visual and cognitive load, visual dilution, placement, target and distractor color, and distractor interference when their insights were transferred from a highly specific and controlled environment to a complex display where all these aspects of attention were required at once. In both experiments, participants searched for a number sequence in an interface that varied to allow testing the effect of the individual attention affordances. In Experiment 1, participants additionally knew the target color, whereas in Experiment 2 they did not. The results for each experiment were discussed in detail above. This section draws together these results and discusses the findings of both experiments with regards to their implication for design research, design application and attention theory.

Implications for Attention Theory

The presented experiments tested predictions from attention theory on the afforded attention allocation of a complex display that mimicked those used in the applied context of a nuclear control room. While the predictions have proven to be robust within the tightly controlled experimental conditions, they should also generalize to real world settings if they truly reflect insights about human nature. In light of this, the experiments presented in this paper thus also have implications for attention theory, as they test of the robustness of the predictions in a complex visual interface.

The most robust finding was that the experiments consistently showed the utility of the seminal theory of Treisman & Gelade (1980) for predicting participants' ability to filter out irrelevant information. In both experiments, results showed that participants were able to filter out the irrelevant graph objects, which indicated that visual clutter was irrelevant for search in complex displays if each of group of objects is distinct from each other, as predicted by the theory.

The second line of findings relates to two theories regarding the effect of visual load on distractor processing: Load Theory and Dilution Theory. Experiment 1, where participants knew the target color, and thus could be distracted by a matching color of the central reactor object, showed that higher visual load had no effect on distractor processing, whereas the presence of irrelevant graphs affected distractor processing in concordance with Dilution Theory (Benoni & Tsal, 2010; Tsal & Benoni, 2010). Despite the larger prominence of Load theory, the findings in the present paper agree better with the more recent Dilution account. The results presented here thus offer some cadence to the claims of Benoni and Tsal (2010; Tsal & Benoni, 2010) and others (Kyllingsbæk et al., 2011; Wilson et al., 2011) that predictions from Load Theory can more accurately be explained by a dilution account of distractor processing. Furthermore, our results corroborate the findings and suggestions of Lleras and colleagues (2017) that Load Theory may have limited generalizability and be of limited applied use. This is further corroborated by the fact that the present paper presents results that conflict with Load Theory's predicted interaction between cognitive load and distractor processing: Experiment 1 showed no interaction effect between distractor processing and visual or cognitive load – in fact, both Experiments 1 and 2 showed faster reaction times under high cognitive load.

Implications for Design Practice

The experiments presented in this paper studied a number of distinct predictions with regards to the afforded allocation of attention to varying parts of a complex display. While the presented display was designed to mimic the interfaces used in a nuclear power plant control room (based on Braseth & Øritsland, 2013), the results should generalize to complex displays in general, given that all findings were related to predicted affordances that were derived from attention theory (which considers attention in a broad manner) and that the manipulated variables were generalizable to other contexts. Notably, the study manipulated cognitive load and visual load, thereby giving design insights pertaining to a wide variety of use cases and complexities. Therefore, the results of this paper contribute several guidelines that designers could utilize to improve complex visual interfaces for both situations.

As the most consistent findings, the experiments showed the importance of reading order in predicting the afforded attention allocation in a complex visual interface. Assuming these results do indeed generalize, high priority items would thus benefit from being presented in the top-left corner of the screen in visual interfaces similar to the ones presented here. Furthermore the experiments consistently showed that additional irrelevant objects were effectively filtered out, and did not afford attention allocation. These results thus indicate that additional cluttering objects do not hinder search in complex visual displays if they are easily distinguishable from the targets users are looking for. In fact, the results of Experiment 1 showed that the irrelevant reactor object has less distracting effect when irrelevant graphs were also present. These results thus indicate, perhaps counter-intuitively, that more visual clutter will reduce the effect of a salient distracting object. As the final consistent finding, both experiments showed, perhaps that higher cognitive load improved speed at a slight cost of accuracy. While further research is needed to elucidate the mechanism behind these findings, the results thus suggest that high cognitive load may be facilitate parts of performance while hindering others while using complex visual interfaces.

On a more specific basis, the experiments showed that the importance of individual color choices in affording attention allocation varies depending on the context: In the high complexity environment of Experiment 1 and 2, the effect of individual colors was only observed when participants had no knowledge of the target color. The results presented here thus represent two different cases for how designers should apply colors in their visual interfaces depending on the expected use case.

Conclusion

This paper presented two visual search experiments that tested predictions from theories of attention in a complex visual interface. It was argued that attention theories can elucidate how attention allocation is afforded in complex displays, and that this can be used to improve how visual interfaces are designed by better aligning the intended and actual use. Furthermore, the experiments acted as a test of whether the predictions of attention theory that have proven robust in the highly controlled and specific displays, also predict the afforded attention allocations in more complex interfaces that involve several aspects of attention at once. Three prominent predictive theories – Feature-Integration Theory, Load Theory, and Dilution Theory – and standalone predictions about attentional guidance and capture of color were tested simultaneously through a display that mimicked both the complex visual interfaces used in nuclear control rooms and the highly controlled displays used in experimental psychology. The results showed that several predictions from attention theories, most notably Feature-Integration Theory, could generalize to predict the afforded attention allocations, and that the accuracy of the predictions related to the amount of knowledge participants had about the target.

Acknowledgements

The research was funded in parts by the Technical University of Denmark, Carnegie Mellon University and the Air Force Office of Scientific Research (AFSOR) under grant No. FA9550-18-1-0088. We furthermore acknowledge partnership with the OECD Halden Reactor Project.

References

Andersen, E., & Maier, A. (2017). The Attentional Capture of Colour in Visual Interface Design : A Controlled-Environment Study. In *Proceedings of the International Conference on Engineering Design, ICED* (Vol. 8, pp. 519–528).

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. https://doi.org/10.1080/13527260500247827

Baars, B. J., & Franklin, S. (2007). An architectural model of conscious and unconscious brain functions: Global Workspace Theory and IDA. *Neural Networks*, *20*(9), 955–961. https://doi.org/10.1016/j.neunet.2007.09.013

Beck, D. M., & Lavie, N. (2005). Look here but ignore what you see: Effects of distractors at fixation. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(3), 592–607. https://doi.org/10.1037/0096-1523.31.3.592

Benoni, H., & Tsal, Y. (2010). Where have we gone wrong? Perceptual load does not affect selective attention. *Vision Research*, *50*(13), 1292–1298. https://doi.org/10.1016/j.visres.2010.04.018

Biggs, A. T., Kreager, R. D., & Davoli, C. C. (2015). Finding a link between guided search and perceptual load theory. *Journal of Cognitive Psychology*, *27*(2), 164–179. https://doi.org/10.1080/20445911.2014.987676

Blijlevens, J., Mugge, R., Ye, P., & Schoormans, J. P. L. (2013). The influence of product exposure on trendiness and aesthetic appraisal. *International Journal of Design*, *7*(1), 55–67. Retrieved from http://search.proquest.com/openview/044b7ee251b3b0c0e1f7a0c670df9f45/1?pq-origsite=gscholar

Boatwright, P., & Cagan, J. (2010). *Built to Love: Creating Products That Captivate Customers*.

Berrett-Koehler Publishers. Retrieved from http://books.google.com/books?hl=nl&lr=&id=POR1qieTPfwC&pgis=1

Boy, G. A. (2017). Human-centered design of complex systems: An experience-based approach. *Design Science*, *3*, e8. https://doi.org/10.1017/dsj.2017.8

Braseth, A. O., & Øritsland, T. A. (2013). Visualizing complex processes on large screen displays: Design principles based on the Information Rich Design concept. *Displays*, *34*(3), 215–222. https://doi.org/10.1016/j.displa.2013.05.002

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, *97*(4), 523–547. https://doi.org/10.1037/0033-295X.97.4.523

Buscher, G., Cutrell, E., & Morris, M. R. (2009). What do you see when you're surfing? In *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09* (p. 21). New York, NY, USA: ACM. https://doi.org/10.1145/1518701.1518705

Buscher, G., Dumais, S. T., & Cutrell, E. (2010). The good, the bad, and the random. In *Proceeding of the 33rd international ACM SIGIR conference on Research and development in information retrieval - SIGIR '10* (p. 42). New York, NY, USA: ACM. https://doi.org/10.1145/1835449.1835459

Cagan, J., & Vogel, C. M. (2001). *Creating Breakthrough Products: Innovation from Product Planning to Program Approval*. FT Press.

Choi, P., Orsborn, S., & Boatwright, P. (2016). Bayesian analysis of color preferences: An application for product and product line design. *Color Research & Application*, *41*(5), 445–456.

Ciavola, B. T., & Gershenson, J. K. (2016). Affordance theory for engineering design. *Research in Engineering Design*, *27*(3), 251–263. https://doi.org/10.1007/s00163-016-0216-5

Crilly, N., Good, D., Matravers, D., & Clarkson, P. J. (2008). Design as communication: exploring the validity and utility of relating intention to interpretation. *Design Studies*, *29*(5), 425–457. https://doi.org/10.1016/j.destud.2008.05.002

Crilly, N., Maier, A., & Clarkson, P. J. (2008). Representing Artefacts as Media: Modelling the Relationship Between Designer Intent and Consumer Experience. *International Journal of Design*, *2*(3), 15–27.

Crilly, N., Moultrie, J., & Clarkson, P. J. (2004). Seeing things: Consumer response to the visual domain in product design. *Design Studies*, *25*(6), 547–577. https://doi.org/10.1016/j.destud.2004.03.001

Crilly, N., Moultrie, J., & Clarkson, P. J. (2009). Shaping things: intended consumer response and the other determinants of product form. *Design Studies*, *30*(3), 224–254. https://doi.org/10.1016/j.destud.2008.08.001

Cutrell, E., & Guan, Z. (2007). What are you looking for? In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07* (p. 407). New York, NY, USA: ACM. https://doi.org/10.1145/1240624.1240690

Dehaene, S., Charles, L., King, J. R., & Marti, S. (2014). Toward a computational theory of conscious processing. *Current Opinion in Neurobiology*, *25*, 76–84. https://doi.org/10.1016/j.conb.2013.12.005

Demir, E., Desmet, P. M. A., & Hekkert, P. (2009). Appraisal Patterns of Emotions in Human-Product Interaction. *International Journal of Design*, *3*(2), 41–51.

Desimone, R., & Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. *Annual Review of Neuroscience*, *18*(1), 193–222. https://doi.org/10.1146/annurev.ne.18.030195.001205

Desmet, P. M. A. (2012). Faces of Product Pleasure: 25 Positive Emotions in Human-Product Interactions. *International Journal of Design*, *6*(2), 1–29.

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some Theoretical Considerations. *Psychological Review*, *70*(1), 80–90.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149. https://doi.org/10.3758/BF03203267

Forster, S., & Lavie, N. (2008). Failures to Ignore Entirely Irrelevant Distractors: The Role of Load. *Journal of Experimental Psychology: Applied*, *14*(1), 73–83. https://doi.org/10.1037/1076-898X.14.1.73

Francis, G., Bias, K., & Shive, J. (2010). The Psychological Four-Color Mapping Problem. *Journal of Experimental Psychology: Applied*, *16*(2), 109–123. https://doi.org/10.1037/a0019095

Gibson, J. J. (1978). The Ecological Approach To the Visual Perception of Pictures. *Leonardo*, *11*(3), 227–235.

Goucher-Lambert, K., & Cagan, J. (2014). The Impact of Sustainability on Consumer Preference Judgments (DETC2014-34739). *ASME 2014 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, *137*(c), 1–11. https://doi.org/10.1115/1.4030271

Goucher-Lambert, K., Moss, J., & Cagan, J. (2017). Inside the Mind: Using Neuroimaging to Understand Moral Product Preference Judgments Involving Sustainability. *Journal of Mechanical Design*, *139*(4), 041103. https://doi.org/10.1115/1.4035859

Hung, W. K., & Chen, L. L. (2012). Effects of novelty and its dimensions on aesthetic preference in product design. *International Journal of Design*, *6*(2), 81–90.

Irwin, D. E., Colcombe, A. M., Kramer, A. F., & Hahn, S. (2000). Attentional and oculomotor capture by onset, luminance and color singletons. *Vision Research*, *40*(10–12), 1443–1458. https://doi.org/10.1016/S0042-6989(00)00030-4

Jameson, K. A., Kaiwi, J. L., & Bamber, D. (2001). Color coding information: Assessing alternative coding systems using independent brightness and hue dimensions. *Journal of Experimental Psychology: Applied*, *7*(2), 112–128. https://doi.org/10.1037/1076-898X.7.2.112

Kahneman, D., & Chajczyk, D. (1983). Tests of the automaticity of reading: Dilution of Stroop effects by color-irrelevant stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, *9*(4), 497–509. https://doi.org/10.1037/0096-1523.9.4.497

Karlsson, M. (2007). Expressions, emotions, and website design. *CoDesign*, *3*(sup1), 75–89. https://doi.org/10.1080/15710880701376802

Kozine, I. (2007). Simulation of human performance in time-pressured scenarios. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, *221*(2), 141–151. https://doi.org/10.1243/1748006XJRR48

Kyllingsbæk, S., Sy, J. L., & Giesbrecht, B. (2011). Understanding the allocation of attention when faced with varying perceptual load in partial report: A computational approach. *Neuropsychologia*, *49*(6), 1487–1497. https://doi.org/10.1016/j.neuropsychologia.2010.11.039

Lau, N., Veland, Ø., Kwok, J., Jamieson, G. A., Burns, C. M., Braseth, A. O., & Welch, R. (2008). Ecological interface design in the nuclear domain: An application to the secondary subsystems of a boiling water reactor plant simulator. *IEEE Transactions on Nuclear Science*, *55*(6), 3579–3596. https://doi.org/10.1109/TNS.2008.2005979

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences*, *9*(2), 75–82. https://doi.org/10.1016/j.tics.2004.12.004

Lavie, N. (2006). The role of perceptual load in visual awareness. *Brain Research*, *1080*(1), 91–100. https://doi.org/10.1016/j.brainres.2005.10.023

Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current Directions in Psychological Science*, *19*(3), 143–148. https://doi.org/10.1177/0963721410370295

Lavie, N., Hirst, A., De Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*(3), 339–354. https://doi.org/10.1037/0096-3445.133.3.339

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics*, *56*(2), 183–197. https://doi.org/10.3758/BF03213897

Leder, H., Carbon, C.-C., & Kreuzbauer, R. (2007). Product-Design Perception and Brand Strength. *Thexis*, *24*(2), 4–7. https://doi.org/10.1007/BF03249147

Lleras, A., Chu, H., & Buetti, S. (2017). Can we “Apply” the findings of Forster and Lavie (2008)? On the generalizability of attentional capture effects under varying levels of perceptual load. *Journal of Experimental Psychology: Applied*, *23*(2), 158–179. https://doi.org/10.1037/xap0000116

Lugo, J. E., Schmiedeler, J. P., Batill, S. M., & Carlson, L. (2016). Relationship Between Product Aesthetic Subject Preference and Quantified Gestalt Principles in Automobile Wheel Rims. *Journal of Mechanical Design*, *138*(5), 051101. https://doi.org/10.1115/1.4032775

Maier, J. R. A., Fadel, G. M., & Battisto, D. G. (2009). An affordance-based approach to architectural theory, design, and practice. *Design Studies*, *30*(4), 393–414. https://doi.org/10.1016/j.destud.2009.01.002

McCarley, J. S., & Steelman, K. S. (2013). Visual attention and display design. In *Neuroergonomics: A Cognitive Neuroscience Approach to Human Factors and Ergonomics* (pp. 51–68). Palgrave Macmillan, London. https://doi.org/10.1057/9781137316523_3

McDougall, S., Tyrer, V., & Folkard, S. (2006). Searching for signs, symbols, and icons: Effects of time of day, visual vomplexity, and grouping. *Journal of Experimental Psychology: Applied*, *12*(2), 118–128. https://doi.org/10.1037/1076-898X.12.2.118

Müller, H. J., Geyer, T., Zehetleitner, M., & Krummenacher, J. (2009). Attentional Capture by Salient Color Singleton Distractors Is Modulated by Top-Down Dimensional Set. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), 1–16. https://doi.org/10.1037/0096-1523.35.1.1

Na, N., & Suk, H. J. (2014). The emotional characteristics of white for applications of product color design. *International Journal of Design*, *8*(2), 61–70. https://doi.org/https://doi.org/10.1007/978-3-319-04798-0_20

Nørager, R. (2009). *Low Level Cognition in User Interfaces*. Department of Psychology, University of Aarhus.

Nordfang, M., Dyrholm, M., & Bundesen, C. (2013). Identifying bottom-up and top-down components of attentional weight by experimental analysis and computational modeling. *Journal of Experimental Psychology: General*, *142*(2), 510–535. https://doi.org/10.1037/a0029631

Norman, D. A. (1988). The design of everyday things. *Doubled Currency*, xxi+257. Retrieved from http://opc-ubm.oclc.org:8180/DB=1/TTL=1/LNG=EN/PPN?PPN=250439824

Orsborn, S., Cagan, J., & Boatwright, P. (2009). Quantifying Aesthetic Form Preference in a Utility Function. *Journal of Mechanical Design*, *131*(6), 061001. https://doi.org/10.1115/1.3116260

Ranscombe, C., Hicks, B., Mullineux, G., & Singh, B. (2012). Visually decomposing vehicle images: Exploring the influence of different aesthetic features on consumer perception of brand. *Design Studies*, *33*(4), 319–341. https://doi.org/10.1016/j.destud.2011.06.006

Reid, T. N., Gonzalez, R. D., & Papalambros, P. Y. (2010). Quantification of Perceived Environmental Friendliness for Vehicle Silhouette Design. *Journal of Mechanical Design*, *132*(10), 101010. https://doi.org/10.1115/1.4002290

Shive, J., & Francis, G. (2013). Choosing colors for map display icons using models of visual search. *Human Factors*, *55*(2), 373–396. https://doi.org/10.1177/0018720812459341

Spence, I., & Efendov, A. (2001). Target detection in scientific visualization. *Journal of Experimental Psychology: Applied*, *7*(1), 13.

Spence, I., Kutlesa, N., & Rose, D. L. (1999). Using Color to Code Quantity in Spatial Displays. *Journal of Experimental Psychology: Applied*, *5*(4), 393–412. https://doi.org/10.1037/1076-898X.5.4.393

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, *74*(11), 1–29. https://doi.org/10.1037/h0093759

Starke, S. D., & Baber, C. (2018). The effect of four user interface concepts on visual scan pattern similarity and information foraging in a complex decision making task. *Applied Ergonomics*, *70*, 6–17. https://doi.org/10.1016/j.apergo.2018.01.010

Stefanucci, J. K., Creem-Regehr, S. H., Thompson, W. B., Lessard, D. A., & Geuss, M. N. (2015). Evaluating the accuracy of size perception on screen-based displays: Displayed objects appear smaller than real objects. *Journal of Experimental Psychology: Applied*, *21*(3), 215–223. https://doi.org/10.1037/xap0000051

Sternberg, S. (1966). High-Speed Scanning in Human Memory. *Science*, *153*(3736), 652–654. Retrieved from http://www.jstor.org/stable/1719418

Still, J. D., & Dark, V. J. (2013). Cognitively describing and designing affordances. *Design Studies*, *34*(3), 285–301. https://doi.org/10.1016/j.destud.2012.11.005

Theeuwes, J. (1992). Perceptual selectivy for color and form. *Percept Psychophys*, *51*(6), 599–606.

Theeuwes, J. (1994). Stimulus-Driven Capture and Attentional Set: Selective Search for Color and Visual Abrupt Onsets. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(4), 799–806. https://doi.org/10.1037/0096-1523.20.4.799

Theeuwes, J. (2010). Top–down and bottom–up control of visual selection. *Acta Psychologica*, *135*(2), 77–99. https://doi.org/10.1016/j.actpsy.2010.02.006

Tovares, N., Boatwright, P., & Cagan, J. (2014). Experiential Conjoint Analysis: An Experience-Based Method for Eliciting, Capturing, and Modeling Consumer Preference. *Journal of Mechanical Design*, *136*(10), 101404. https://doi.org/10.1115/1.4027985

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Tsal, Y., & Benoni, H. (2010). Diluting the Burden of Load: Perceptual Load Effects Are Simply Dilution Effects. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6), 1645–1656. https://doi.org/10.1037/a0018172

Vazquez, E., Gevers, T., Lucassen, M., van de Weijer, J., & Baldrich, R. (2010). Saliency of color image derivatives: a comparison between computational models and human perception. *Journal of the Optical Society of America A*, *27*(3), 613. https://doi.org/10.1364/JOSAA.27.000613

Wagman, J. B., & Carello, C. (2003). Haptically Creating Affordances: The User-Tool Interface. *Journal of Experimental Psychology: Applied*, *9*(3), 175–186. https://doi.org/10.1037/1076-898X.9.3.175

Ware, C. (2008). *Visual Thinking: For Design*. *Ergonomics* (Vol. 53). Morgan Kaufmann. https://doi.org/10.1080/00140130903458285

Weyer, U., Braseth, A. O., Eikås, M., Hurlen, L., Kristiansen, P., & Kvalem, J. (2010). Safety Presentation in Large Screen Displays - A New Approach. In *SPE Intelligent Energy Conference and Exhibition*. Society of Petroleum Engineers. https://doi.org/10.2118/128666-MS

Wilson, D. E., Muroi, M., & MacLeod, C. M. (2011). Dilution, Not Load, Affects Distractor Processing. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(2), 319–335. https://doi.org/10.1037/a0021433

Wolfe, J. M. (2007). Guided search 4.0. *Integrated Models of Cognitive Systems*, (3), 99–120. https://doi.org/10.1007/978-94-011-5698-1_30

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 419–433.

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, *5*(6), 495–501. https://doi.org/10.1038/nrn1411

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*(3), 58. https://doi.org/10.1038/s41562-017-0058

Xenakis, I., & Arnellos, A. (2013). The relation between interaction aesthetics and affordances. *Design Studies*, *34*(1), 57–73. https://doi.org/10.1016/j.destud.2012.05.004

Yamani, Y., & McCarley, J. S. (2010). Visual search asymmetries within color-coded and intensity-coded displays. *Journal of Experimental Psychology: Applied*, *16*(2), 124–132. https://doi.org/10.1037/a0019570

You, H. chen, & Chen, K. (2007). Applications of affordance and semantics in product design. *Design Studies*, *28*(1), 23–38. https://doi.org/10.1016/j.destud.2006.07.002

CHAPTER 4:

Biased Information Search in Emergencies

4.1. Chapter Introduction

Having discussed attention, the primary underlying mechanism that governs what information enters consciousness, and the application for improving visual design for emergencies in the previous chapter, this chapter focuses on the higher-level decision making that determines where information is searched for and, when it is found, how it is interpreted. While the attention system will ultimately be responsible for the search and discovery of the required information, it is the decision making process of e.g. operators that determines which documents and/or screens are presented to the attention system. Furthermore, once the information is found, the interpretation of the information will influence what actions will be performed as a consequence. Therefore, knowledge about the decision making processes that underlie information search and interpretation is equally valuable for improving visual design for emergencies.

In the traditional economic model of decision makers, the decision making process behind finding and interpreting information would be considered to be rational process where logic would be the main driver behind figuring where to go next (Tversky & Kahneman, 1981). However, the last five decades of research has seen a large accumulation of evidence that has challenged this notion of the rationality, instead demonstrating that human decision making is guided by a series of biases and heuristics that consistently, and predictably, skew their choices and judgments. The most popular explanation for this deviation from rationality comes from Dual Process theories (e.g. Evans, 2003; Kahneman & Frederick, 2002), which have proposed that human decision making relies on two systems rather than one: System 1, which is fast, non-conscious, and prone to biases, and System 2, which is slow, conscious and (usually with great effort) able to produce rational responses (Evans, 2003, 2008). We are thus able to produce rational and logical responses, however, due to the brain's preference towards conserving energy (Kahneman, 2011; Kool et al., 2010), we inadvertently tend to base much of our decision making on System 1 thinking, which in turn leads to risks of biases.

Operators of nuclear power plant control rooms, as well as similar areas such as medicine, aviation and military, are usually highly trained experts within their domain. Based on this expertise, research has shown that they are likely to employ highly refined heuristics that make their decision making processing faster for the domain they are trained in (Dreyfus, 2004; Gigerenzer & Gaissmaier, 2011; Shanteau, 1992). However, when conditions are non-typical, as is often the case for emergencies within the nuclear realm (U.S. Nuclear Regulatory Commission, 2002), these heuristics may lead to sub-optimal behaviour (Kahneman & Klein, 2009). Furthermore, emergencies are often associated with high cognitive load (Burian, 2005), which increases the risk of biased decision making (Gailliot et al., 2007; Muraven & Baumeister, 2000; Muraven et al., 1998). Taken together, investigating whether experts are susceptible to decision making biases in emergencies, and, if yes, which ones, is critical for preventing adverse events. Given that operators rely on visual aids, most notably written procedures, designers may be able to prevent or minimize biases through design interventions (Cook & Smallman, 2008; Walmsley & Gilbey, 2017). Conversely, the visual aids may worsen the likelihood of biased behaviour if designed improperly.

Based on this, the research presented in this chapter is concerned with detecting whether nuclear control room operators display biased decision making, whether any such biases are amplified with the procedures that they rely upon, and with suggesting ways in which designers and design researchers could look to prevent such biases. The research was conducted in collaboration with researchers at the Institute for Energy Technology in Halden, Norway, and consists of literature review and studies of decision making of real-life operators in realistic scenarios presented in the Halden Man-Machine Laboratory (HAMMLAB).

4.2. Research and Findings

The paper presented in this chapter, (Paper 6/ Andersen et al., 2018), is a conference contribution wherein it was investigated whether operators display confirmation bias and biased decision making due to misapplied expertise. Furthermore, it was investigated whether the operating procedures could have increased the likelihood of these biases occurring. The study was done through a re-analysis of two previously conducted simulation experiments in HAMMLAB to investigate the decision making of crews of active, experienced nuclear control room operators in scenarios that required them to step outside of procedural guidance to avoid adverse effects. The re-analysis was conducted by analysing the behaviour of operators using a detailed description of performance created by a nuclear control room expert, as well as by revisiting audio and video recording of the experiments. It was found that the majority of crews showed evidence of biased behaviour due to both confirmation bias and misapplication expertise, and that these biases (particularly confirmation bias) were amplified by the presence of the procedures. Based on this, suggestions were made with regards to future studies as well as design interventions that could alleviate the biases rather than aggravate them. The results thus both have implications for understanding expert decision making in emergencies and for preventing decision making biases through improving the design of operating procedures.

4.3. Implications for Design Practice

Besides providing knowledge of operator behaviour, which may inform designer decisions, the results have direct implications for the design of procedures for emergencies. The findings that procedures could amplify the likelihood that the operators were biased indicate that procedure designs should be explored to counter-act this effect. The most successful method for combating confirmation bias, which our results indicated that procedures may have aggravated, is to incorporate counterfactuals or prompts for considering the opposite of the current hypothesis (Cook & Smallman, 2008; Galinsky & Moskowitz, 2000). The results of Paper 6 showed potential evidence that this could be an effective intervention to be used in design of procedures for nuclear control rooms, as the crew least likely to be affected by confirmation bias was readily applying counterfactuals in their communication strategy. The results thus support practice by both pointing out a potential cause for variation in performance and behavioural patterns of active, experienced nuclear control room operators, and by offering suggestions for what could be an effective intervention for procedure design, which could better aid visual search strategies of operators.

Paper 6:

Biased Decision Making in Realistic Extra-Procedural Nuclear Control Room Scenarios (2018), by Andersen, Kozine, I., and Maier, A. In *8th International Conference on Design Computing and Cognition*

Biased Decision Making in Realistic Extra-Procedural Nuclear Control Room Scenarios

Emil Andersen, Igor Kozine and Anja Maier
Technical University of Denmark, Denmark

In normal operations and emergency situations, operators of nuclear control rooms rely on procedures to guide their decision making. However, in emergency situations these procedures may be insufficient in guiding operators. Little is known about the decision making strategies that operators employ in these extra-procedural situations. To address this, a realistic simulation study was conducted with five crews of active, licensed nuclear operators to see the behavioural patterns that occur when procedures are not sufficient. This paper, a re-analysis of a previously collected dataset, investigates how the design and existence of procedures influence, and possibly bias, decision making. Evidence is found that operators were affected by confirmation bias, and that mismatches between their home power plant and the simulated power plant made them commit errors due to misapplied expertise. Furthermore, this effect was amplified by the existence and design of the procedures used. Avenues for debiasing through design are discussed.

1. Introduction

Studies of operators in nuclear control rooms, airplane cockpits and medical decision making have led to a greater understanding on decision making in high-stakes complex environments over the last several decades. To deal with the complex requirements of these environments and strengthen performance, researchers in decision making and design have worked hand in hand to improve the interfaces, environments and tools of the specialists that operate these fields (e.g. [1]–[5]). An important early development was the move towards the use of written procedures and checklists. Usually in the form of physical paper copies of varying length, procedures and checklists are written documents that specify conditions for their use, followed by a list of diagnosis and action steps. However, previ-

Design Computing and Cognition DCC'18. J.S. Gero (ed), pp. 601-620.

ous studies have found that for the critical situations in nuclear control rooms, the majority of real life operating events included non-typical conditions [6], [7]. In such situations, where the predicted situation in the procedure does not match the observed situation, procedures may become inefficient or lack proper guidance [8].

In the design field, an increasing amount of studies have sought to uncover patterns in human behaviour in order to guide designers in their efforts, e.g. on topics ranging from basic perceptual functioning, to aesthetic product preferences and other critical factors for consumer decisions (e.g.[9]–[17]. Through evaluations of real and/or stylized products and product attributes, these studies have started to shape our understanding of how humans in general and in relation to specific groups in particular perceive and interact with various design characteristics. Furthermore, studies of designers have found biases in their decision making such as design fixation ([18], for a recent review see [19], the preference effect [20], strategies for how these effects can be mitigated, and how these strategies interact with expertise [21]. Outside of the design field, developments over the last several decades have led to an increased understanding of decision making strategies and biases in general [22], [23]. However, little is known about how these decision making insights apply in the practical situation of a control room emergency, and whether the decision making biases are reduced or amplified by the existence of the designed objects such as procedures and checklists – particularly in the non-typical situations that characterise real life emergencies.

To address this, the present paper re-analyses data collected for a project involving two realistic pressurized water reactor scenarios conducted at the Halden Man Machine Laboratory (HAMMLAB) simulator in 2014. The scenarios were designed such that multiple complications would lead to situations where crews had to perform autonomous extra-procedural actions to achieve optimal performance. The results of the original study were documented by Massaiu & Holmgren [24]. They investigated how operators perceive discrepancies between their own plans and the procedure, how crews compromised between needing to act fast and to follow procedures, and how the crew size and composition affected diagnosis and decision making. They found that crews, with some exceptions, prioritised strict adherence to procedures and that crew size and composition did not influence performance. The scenarios were described in detail by Massaiu & Holmgren [25] to allow for future re-analysis, such as this paper.

Adding to this former work, the present paper aims to show biases and heuristics that may have caused divergences in behaviour amongst the crews. The purpose of this study then is to create an exploratory platform to show how biases may influence expert decision making in critical situations, as well as to begin shaping our knowledge of how these biases may arise from the designed objects that these operators interact with during their work.

2. Background

In this paper, we focus on two biases that have been related to expert decision making: The first is the bias that occurs when expertise is transferred to a similar but different situation, thus causing misapplications of one's expertise. The second is confirmation bias, which is the tendency to overly prioritise and seek for information that benefits existing views. Both biases have been shown to impact decision making of experts in many diverse fields, such as medicine, engineering and law [26].

In this section, these biases are described, and we outline which behavioural patterns should be observed if the nuclear control room operators were affected by them.

2.1. Expertise

2.1.1. What is an Expert?

In this paper, we use as the basis for our definition of expertise the one given by Simon [27]: *"The situation has provided a cue: This cue has given the expert access to information stored in memory, and the information provides the answer. Intuition is nothing more and nothing less than recognition"*. From this viewpoint, an expert is one who has been exposed to a high variety of situations and has learned the correct response, which allows him to swiftly recall and apply it in future situations. Furthermore, we extend our definition of expertise based on the arguments by Kahneman & Klein [22]. Drawing on a study by Shanteu [28], they argue that expertise will only form if a) the context of training provides valid cues for learning, meaning cues that reflect real patterns in the context, and b) if the context of training is sufficiently regular to allow for learning of patterns. Without these aspects, they argue, it is not possible to learn whether your behaviour is resulting in good or bad results, and thus expertise cannot be achieved. The nuclear power plant is a vastly complex system, with a myriad of technical details that needs to be acquired over several years an operator is certified. However, from the criteria of Kahneman & Klein [22], the nuclear power plant control room is a valid context for acquiring expertise, as the relation between inputs and outputs of actions is both consistent and readily available of observation for the operators. Given the extensive training required for certification, as well as the substantial experience of all the participants in the study, a high level of expertise should thus be expected.

2.1.2. Misapplied Expertise

Given this expertise, it is expected that the operators employ highly refined heuristics (shortcuts for decision making) that allow them to make (near-) optimal decisions for the context they have been trained in, with lower effort [23], [29], [30]. However, these strategies can decrease performance if applied to other contexts, where they are not adequate. For the present simulation study, this may be the case. First, the crews were trained at a power plant in a different country than the simulated power plant. Second, not all parts of the simulation perfectly matched what would be observed in the reference power plant. The operators could perform sub-optimally due to lack of plant-specific knowledge or due to expectations of and/or reliance on signals that do not come due to plant differences. Furthermore, due to these differences, the operators need to adjust their behaviour to reflect a lower level of expertise than what they have for their home plant, taking the more explorative mind-set of a novice. However, previous research has shown that experts, when put in a similar but not identical situation to what they have expertise for, tend to act as if their expertise also applies to the novel situation [22]. This is caused by a false belief that there is a perfect transfer of skill between the two situations. While the operators receive training in operating the simulation power plant prior to the simulation scenarios, there may nevertheless be deviations between the two power plants that will cause operators to use heuristics that are inappropriate for the specific context.

If the nuclear power plant operators are affected by the bias of misapplied expertise, we expect that one, they will double-check their decisions in an insufficient manner, as they would not need to do this if they were highly trained, and two, they will deviate in ways that turn out wrong due to lacking plant-specific knowledge.

2.2. Confirmation Bias

Confirmation bias is the non-conscious tendency to seek for- and to give higher value to information that confirms our existing views and, conversely, to ignore and deprioritise information that goes against our existing views. Confirmation bias is thus an overarching term that covers the tendency to strongly persist in existing beliefs, as a result of biased evaluation of information and in search for information [26].

2.2.1. Belief Persistence

The first aspect of confirmation bias is belief persistence, which is the term for a collection of tendencies that cause early beliefs to be very resistive to change: First, the tendency to persist in early hypotheses for no reason than them being the first adopted hypotheses [31]. Second, the tendency to be more likely to question information that contradicts their existing belief, while being less likely to question information that confirms their pre-existing belief [32], [33]. Third, the tendency to be likely to explain away events as random etc. if they conflict with their existing beliefs, thus discrediting the events rather revisiting the belief [34].

If operators are susceptible to belief persistence, we expect that operators will persist in their early hypotheses if they do not contradict the operating procedures (regardless of whether or not the procedures are correct for optimal decision making at the time).

2.2.2. Biased Search for Information

The second aspect of confirmation bias is the tendency to only seek information that confirms one's existing view, or to only seek for information that would only exist if the existing view was correct. Conversely, it is the tendency to avoid information that would disconfirm one's view and/or not to seek for information that would exist if an alternate view was correct [35]. This tendency thus allows one to never disconfirm one's view through never exposing oneself to situations that threaten the viewpoint. Furthermore, given that one only samples information that supports the view, confidence in the view increases [36].

Similarly, we expect that operators will perform confirmatory search by looking at power plant locations that will show problems only if their hypothesis was true and will tend not to search for disconfirming information through, e.g., alternate sources such as field operators.

3. Case Study

The data that form the basis for this paper are the decisions of nuclear control room operators in two realistic simulation scenarios conducted in the HAMMLAB simulator in 2014. Two scenarios of realistic emergency situations in a Pressurized Water Reactor were run by five crews of 3-5 crew members. The size of the crews and the exact scenario durations are shown in Table 1. In nuclear operations, operators rely on Emergency Operating Procedures to solve emergencies. The operators have knowledge of a vast array of 'entering conditions' for various procedures, and will 'enter' a given procedure in response to these conditions. Once entered, the procedures will guide the operators through identifying and alleviating problematic symptoms. Operators are at no point required to know the cause of the observed problems, only to follow the procedures that instruct how to respond to these symptoms.

Table 1: Crew size and duration of scenarios. Note that shorter duration does not necessarily indicate better performance.

Crew	1	2	3	4	5
Size	5	5	3	4	3
Scenario 1 Duration	02:06:21	01:54:47	02:50:00	01:52:51*	02:10:37
Scenario 2 Duration	01:26:21	01:14:47	02:10:00	01:12:51*	01:30:37

*Scenario was stopped before the crew had completed the final goal

The two scenarios are unique with respect to the cause of the problem. However, in both scenarios, emergency operating procedures are entered in response to the reactor 'tripping' (this term refers to neutron absorbing control rods being inserted into the core, thus stopping the chain reaction). While tripping the reactor stops further power from being produced, the power plant is not safe until problems such as leaks causing spread of radiation are solved, and the plant is cooled and depressurized. Until safe shutdown is achieved, adverse effects such as release of radioactive material to the atmosphere, or, in the worst case, core meltdown, are still possible.

Both scenarios were designed such that following operating procedures was not sufficient for safe and effective shutdown. Operators were thus required to perform autonomous actions to avoid adverse effects. The scenarios are described in detail below. Overall, the complex problems of the scenarios caused several problems for all crews in both scenarios, albeit to varying degrees for the various crews, as will be elaborated below.

3.1. Scenario 1

The first scenario involved multiple leaks on the piping system that connects the core with the plant's three steam generators. The crew is given the cover story that construction is ongoing nearby, which, shortly after start, is cited as the cause for a blast that gives vibrations to the plant, including the control room. In the simulation, this blast results in immediate release of radioactive material in all steam generators, followed shortly (12 mins after start) by a small leak in a tube connected to steam generator #2, and subsequently (20 mins after start) a rupture in a tube in steam generator #3. The leak in steam generator#3 will increase in size two times, first at 25mins and then at 40mins after start, with the latter being equivalent of a complete tube rupture. If the crew has not manually tripped the reactor at 40mins, the automatic tripping system will do so shortly after the complete tube rupture.

The challenge for the crew is ensure cooldown while avoiding using the two damaged steam generators' relief valves, as this would result in release of radioactive material to the atmosphere. To do so, the two damaged steam generators should be isolated. This task is complicated by the fact that it is not clear from following the procedures and the information displayed whether steam generator 2 is causing problems, as it is obscured by the effects of the rupture in steam generator 3. Operators must thus actively look for additional information to successfully handle the task.

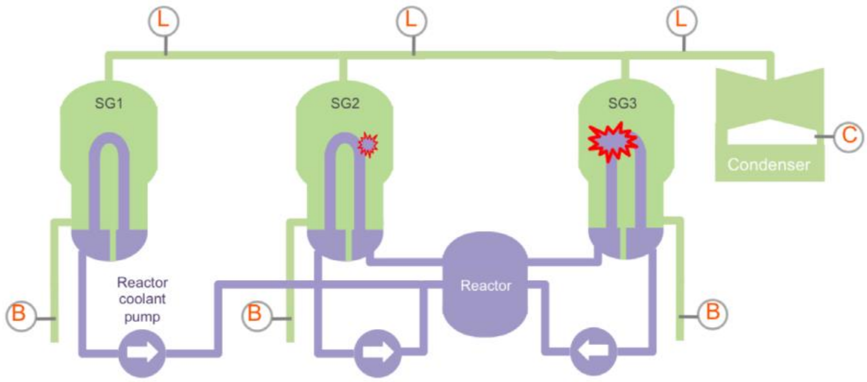


Figure 1: Diagram of Scenario 1, courtesy of Massaiu & Holmgren [24].

3.2. Scenario 2

The second scenario involves an irreversible loss of coolant following leaks to the Reactor Coolant System, which results in water spilling on the floor of the auxiliary building. The scenario begins with a distracting task in the form of a pump trip. This will occupy the operators at the start of the scenario. The first major complication happens when two valves starts leaking in the Residual Heat Removal System (one at start, the other after 8 minutes). At around 11 mins from start, a pipe in the Residual Heat Removal system of the auxiliary building will break, resulting in reactor coolant fluid spilling on the floor. Finally, a smaller leak will occur in the Reactor Coolant Pump thermal barrier, which will complicate the detection of the primary leak. The loss of pressure will cause an automatic trip of the reactor if it is not initiated manually.

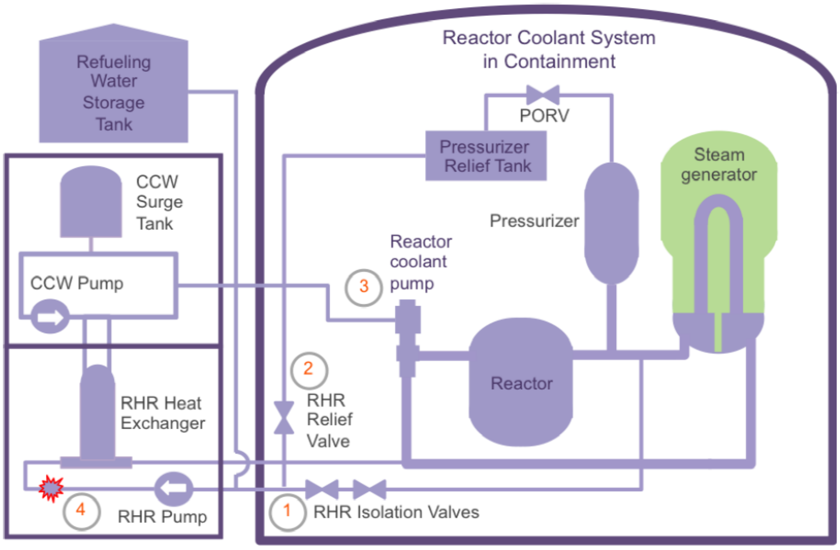


Figure 2: Diagram of Scenario 2, courtesy of Massaiu & Holmgren [24].

The challenge for the crew is to ensure safe and effective cooldown while reducing the effects of the leaks. This task is complicated by the fact that the procedure for this type of event aims at identifying the main leak and to isolate it, but in the present scenario the procedure's directions are insufficient. Furthermore, although the operators do not know that the leaks are not isolable, it is considered optimal performance to discover the location of the leaks and to try isolation actions, beyond the procedures' guidance.

4. Methods

4.1. Dataset

Five crews of certified operators from three nuclear power plants were recruited. Two crews had three members, two crews had four members and one crew had five members. All crews participated in both scenarios.

The study took place at the Halden Man Machine Laboratory (HAMMLAB) at the Institute for Energy Technology, Halden, Norway, in a realistic simulation set-up that mimics a Swedish Pressurized Water Reactor (PWR). Audio and video materials were recorded during the scenarios, which serve as the raw dataset for this analysis. The audio material included all conversation between operators, sounds played in the environment, conversations between the operators and the experimenters (who, at various times, roleplayed as field operators) and conversations between experimenters. The video material consisted of four streams of serially played still-shots of the operators and recordings of the displays used by the operators (including mouse movements on these). The raw dataset was processed by a former shift supervisor from the simulated plant, who has many years of experience as an operator and thus has the required skills to evaluate performance, together with a human factors specialist from the Halden Reactor Project. These evaluations and additional comments were noted down in detail in a spreadsheet, along with the timestamp and a brief description of the various events that the process expert had deemed significant. The contents of the comments included, but were not limited to: examples of good behaviours, errors and deviations from protocols. To measure performance, each team was evaluated by the process expert using a 5-points rating scale, which gives a score of one to five on task critical operations, and with five being optimal performance. For the purpose of this study, the crew performance scores for each scenario were averaged as an indicator of overall performance.

As this study re-analyses previous data, the dataset for this paper includes the aforementioned raw data and processed spreadsheet, as well as a detailed report with further processed descriptions of the events [25]. The study was approved by the Ethical Council of the Institute of Energy Technology and complied with their ethical guidelines.

4.2. Analysis Procedure

To perform the analysis for this report, the first author first read the report by Massaiu and Holmgren [25], which gives a detailed description of the scenarios for each crew. This was to better familiarise himself with the terminology and to look for possible connections to theory. Second, he analysed the detailed spreadsheets, marking events that were of particular interest for further processing. Third and final, he re-inspected the video and audio material to find examples of specific dialogue exchanges, and to listen for cues, such as tonality, formality of word usage, volume in the speech, overall to get a better understanding of the marked events. In order to ensure that the context was properly understood, the procedure for re-inspecting an event was to start the recording approximately one minute before the timestamp of the event and to end approximately one minute after the event had ended. The purpose of the re-inspection was to ensure that the writing of the process expert (which did not contain comments about tonality etc.) was not misunderstood. This was especially important when the process expert had written ad verbatim quotes from the operator dialogue.

Following the re-inspection of the data, the first author noted all examples of good behaviour (as commented by the process expert) in the aforementioned spreadsheet with a green mark and all examples of errors and deviations with a red mark. Following this, these behaviour examples were coded into categories and the frequency of each category was counted to give an overview of the performances.

Based on these steps, behavioural patterns of the highest, medium, and lowest scoring crews as defined by the crew performance ratings were compared.

5. Results

In this section we show the behavioural findings from the simulation scenarios. The results are divided by scenario, and for each scenario we describe the behavioural patterns of the top-, medium- and lowest-scoring crew(s). The differences between these behavioural patterns and the extent to which they are caused by use of procedures and decision making biases will be discussed in section 6.

5.1. Scenario 1

In scenario 1, the major point of divergence in team performance lied in identifying that both steam generator #2 and #3 suffered from structural damage. Two of five crews detected that there were problems with both steam generators, whereas the remaining three crews proceeded as if only one steam generator was leaking at any given time. However, all teams were challenged in this detection, as it adhered to procedure and was plausible that any effects observed from steam generator #2 could have been caused by ‘shine’ (which is the term for radioactive measures spilling over from a larger nearby rupture). The crew performance scores for scenario 1 are found in table 2. Problems were observed for all teams and radioactive material was released to the atmosphere by all crews (crew 5 was the only crew who did so intentionally). As was previously reported by Massaiu & Holmgren [24], crew size did not predict performance and the teams with a Shift Technical Advisor did not outperform teams without a person performing this role.

In the following, we characterise the differences between the top, medium and lowest scoring crews, by describing their behavioural patterns. Differences in these behavioural patterns will be discussed in concordance with our hypotheses in section 6.

5.1.1. Factors Causing High Performance

The highest scoring crew was crew 5, with an average performance score of 4.7. Four aspects were observed that may have caused this performance: First, the process expert observed few technical errors and/or instances of suboptimal execution. Second, the crew was very active in looking for alternate sources of information and in testing multiple hypotheses. For example, crew 5 was the only crew to autonomously ask for information about Steam Generator #2 integrity, follow up to ensure that they received the information and perform actions to isolating Steam Generator #2. The focus on looking for alternate sources and on investigating alternate hypotheses was also visible in the language used in strategy meetings, where crew members would use utterances such as *"ruptured steam generator or generators"* or *"we can't tell WHICH steam generator"* to describe the problem. Third, the Unit Supervisor did not read aloud the notes in the procedures and was, along with the rest of the crew, very proactive in planning future actions. Specifically, the crew frequently called for status updates, wherein ideas were discussed and shared and plans were laid for future actions. Fourth, the Unit Supervisor in crew 5 would read all procedure steps aloud before any actions were taken, whereas in other crews, the Unit Supervisor would read each procedural one step at a time as they were completed.

5.1.2. Factors Causing Medium Performance

The medium-scoring crews were crew 1, 2 and 3, with average performance scores of 3.8, 3.7, and 3.8. Of these, one crew, crew 3, identified that both Steam Generator #2 and #3 had suffered structural damage. However, all three crews considered, to varying degrees, whether Steam Generator #2 had leaked as well. The different strategies for investigating these considerations were as follows: Crew 3 called a 'Field Operator' (roleplayed by the control room), but without opening sampling valves (thus not allowing for sampling), and did not call the field operator back after the sampling valves had been opened. After some deliberation in the experimenters' gallery, the 'Field Operator' decided to call the crew and share the information that both Steam Generator#2 and Steam Generator#3 showed radiation, after which the crew promptly performed actions towards isolating Steam Generator#2 as well. However, the crew used Steam Generator#2 for cooldown. Crew 1 and 2 discussed the possibility of a leak in Steam Generator#2, opened sampling valves and sent a 'Field Operator' to collect samplings. However, both crews did not follow up with the Field Operator and abandoned the hypothesis that Steam Generator#2 could be leaking as well. Compared to crew 5, the three middle scoring crews used singular terms about the steam generator problem early on, such as *"the ruptured steam generator" (crew 1)*. Notably, some medium scoring crews originally believed the measures in Steam Generator#3 were due to shine (when it had first leaked) and then changed their hypothesis to it being the Steam Generator#2 measures that were caused by shine. These crews thus essentially did not change the hypothesis that it was only one steam generator that was faulty – they simply changed their mind about which one it was. Finally, the medium scoring crews tended to search for information that backed up (rather than falsified) their views, as is reflected in language such as *"request chemistry back that up with local sample"(crew1).*

5.1.3. Factors Causing Low Performance

The lowest scoring crew was crew 4, with an average score of 2.7. Two aspects were observed that may have caused this performance: First, the process expert observed technical errors in executing the procedures. The effects of these errors caused ripple effects that made the scenario more and more complex, eventually resulting in the scenario being ended before depressurization was achieved. Second, we observed examples of inefficient communication between crew members. For example, the crew's Reactor Operator and Shift Technical Advisor suggested several times to perform steps to test whether Steam Generator #2 was also leaking. However, the Unit Supervisor was of a different belief, and thus did not translate the recommendations into actions towards isolation. As a result, the crew quickly abandoned the possibility that two Steam Generators were damaged and instead focused on Steam Generator #3.

Table 2: Crew size, duration of scenario and performance scores for Scenario 1.

Scenario 1 Performance Scores					
Crew	1	2	3	4	5
Size	5	5	3	4	3
Alarm Handling	5	4	4	3	5
Identification and Isolation	3	4	3	3	4,5
Cooldown	3	5	4	4	5
Depressurization	4	1	4	3	4,5
Stop Safety Injection	n/a	5	5	2	5
Pressure Balance	n/a	3	2,5	1	4
Average	3.8	3.7	3.8	2.7	4.7
Duration	02:06:21	01:54:47	02:50:00	01:52:51*	02:10:37

*Scenario was stopped before the crew had completed the final goal

5.2. Scenario 2

In scenario 2, the major point of divergence was the degree to which teams chose to invest resources into identifying the leak locations by going outside of the procedures. These differences are detailed below. Overall, the impact of choosing to use resources to identify and isolate the leaks varied: Teams hit procedural goals at comparable speeds, with crew 3 being slightly faster. Using resources for non-procedural operations thus did not slow down the progression through procedures.

5.2.1. Factors Causing High Performance

The highest scoring crew was crew 2 with an average performance score of 4.0. Two aspects were observed that may have caused this performance: First, the crew was one of two crews (along with crew 5) that invested heavily into finding the location of the leaks. The crew decided early on that they needed information from external sourced information. Therefore, they communicated frequently with 'Field Operators' (roleplayed by the Control Room) throughout the scenario. Based on these communications, they were able to identify both leaks and to perform actions towards isolating them. Second, the crew's Unit Supervisor chose not to read notes in the procedures aloud. In addition to these performance measures, the crew was furthermore the only crew to execute on restoring water to the Refuelling Water Storage Tank.

5.2.2. Factors Causing Medium Performance

The medium scoring crew was crew 1, with an average performance score of 3.3. One aspect was observed that may have caused this performance: Compared to two of the lower scoring crews, crew 4 discussed and performed some preliminary actions towards identifying and isolating the leak. However, rather than spending resources on investigating further the exact locations, they tried to deduce the locations from secondary information instead of testing their hypotheses with alternative sources of information, whereas crews 2 and 5 relied on dialogue with 'Field Operators'. Based on this information gathering, crew 1 was able to achieve some degree of identification of the location of the leaks, and to perform some isolating actions.

5.2.3. Factors Causing Low Performance

The lowest scoring crews were crew 3, 4 and 5, with average performance scores of 2.3, 2.0 and 2.5 respectively. Different aspects were observed for each team that may have caused this lower performance: For crew 3 and 4, the lower score was caused by the fact that they did not attempt to identify and isolate the leaks. For crew 5, their lower score was caused by suboptimal performance during attempts at isolating the leak in the Reactor Coolant System and during cooldown.

Crew 3 and 4 had preliminary suspicions and discussions about the possibility of a leak. Crew 3 decided early on that there was only one minor leak and to not spend resources on communicating with a Field Operator regarding alternative information – they only communicated with the Field Operator for practical tasks, such as energizing valves. Crew 4 discussed calling a FO to investigate further, but decided not to do so as they believed there to be too many possible candidates for the leak location.

A common factor for the lower scoring crews was that they did not show appropriate patience in watching the effects of the procedure actions. As a result, the crews entered subsequent procedures based on misleading information about whether or not the preceding procedure had been effective. Crew 5 attempted to compensate for this by re-running procedures while simultaneously entering another procedure. This increased the workload on the crew, which may in turn have caused the lower performance with regards to cooling.

As a final factor, crew 5 was, as in scenario 1, very active in calling for strategy briefs and in collaborative planning of future steps.

Table 3: Crew size, duration of scenario and performance scores for Scenario 2.

Scenario 2 Performance Scores					
Crew	1	2	3	4	5
Size	5	5	3	4	3
Attempt to identify?	Limited	Yes	No	No	Yes
Identification of leak in RCS	3	4	1,5	1,5	2
Identification of leak in RHR	3	4	1,5	1,5	4,5
Cooldown	4	4	4	3	1
Average	3.3	4.0	2.3	2.0	2.5
Duration	01:26:21	01:14:47	02:10:00	01:12:51*	01:30:37

*Scenario was stopped before the crew had completed the final goal

6. Discussion

This study investigated the decision making and performance of five crews of nuclear control room operators in two realistic simulated scenarios. The scenarios involved non-typical situations, which were caused by multiple failures in the nuclear power plant system and could only be detected in full through deviation from procedures, such as autonomous requests or search for additional information. We found that some crews strictly adhered to procedures despite this leading to suboptimal performance. Furthermore, we found that crews did generally not persist in the mind-set of testing multiple hypotheses once procedures had been entered. This was seen in the reformulation and/or further commitment to the hypothesis that only a single steam generator was leaking, which was consistent with the procedure in Scenario 1 and a rationalisation/explaining away of reasons to go outside of procedures in Scenario 2. Crucially, these behavioural patterns were observed despite the fact that, generally, teams who also pursued actions outside of procedures were not slower or less accurate. Furthermore, we found examples of both types of bias behaviour. In the following, we elaborate our findings with respect to each bias and suggest implications for design.

6.1. Expertise

Despite crews receiving extensive training in the simulated power plant, we observed several possible indicators of biased behaviour due to misapplied expertise: First, in scenario 2, the decision of crew 1, 3 and 4 to not allocate additional resources towards detecting the leaks could have been an optimal strategy at their home plant. Their expertise may have thus guided them not to continue, as conserving resources would lead to greater success. Second, in scenario 2, crew 1 chose not to collect additional information from outside sources as they believed they could reach sufficient data from secondary calculations in the control room, and crew 4 decided not to pursue identification because they believed there to be too many possible causes. These behavioural patterns are consistent with the tendency not to seek for additional information due to expertise, which, in this case, caused suboptimal performance.

6.2. Confirmation Bias

6.2.1. Early Hypotheses

Despite not being required to do so by the procedures, we observed that all crews created hypotheses about the cause of the problems early on in both scenarios. Furthermore, in both scenarios, we found that for the majority of crews, the first hypotheses persisted through the entire scenario. Deviations were seen in scenario 1, where crew 5 continuously explored multiple hypotheses throughout the scenarios (thus never committing to just one hypothesis) and where crew 3 committed to a hypothesis but reconsidered when salient outside information from a field operator caused them to reconsider. In particular, as evidence by the performance of crew 4 in scenario 1, it seemed that the early beliefs of the Unit Supervisor were especially important for choice of strategy. These observations are thus consistent with the notion that operators were affected by confirmation bias in the form of commitment to early hypotheses.

6.2.2. Confirmatory Search

As expected, we found several examples of confirmatory search. In scenario 1, all but one crew had adopted the hypothesis, which was consistent with the procedures, that only one Steam Generator was damaged. To test this hypothesis in a non-confirmatory manner, crews would need to look for information about the integrity of the other steam generators and see whether there was damage in multiple locations. Furthermore, given that there was a considerable delay between Steam Generator #2 and #3 leaking, the hypothesis that only a single Steam Generator was damaged was true for an extended period of time. To avoid confirmatory search, crews would thus have to continuously look for changes in information that caused the previously true hypothesis to be false. We observed that only one crew, crew 5, employed a strategy that allowed them to continuously search for alternate sources of information, while another crew, crew 3, was given alternate information by the field operators, which prompted them to adopt the true hypothesis that two leaks had occurred. For the remaining crews, two sources of information were used to confirm the original hypothesis, which caused the teams not to search for additional information. First, after the rupture in Steam Generator#3 had reached its maximum, the severity of its effects was much larger. Consequently, crews could readily explain radiation measures from Steam Generator #1 and Steam Generator#2 as 'shine' effects which is a common occurrence and thus a theoretically valid data point to confirm that there was only damage to Steam Generator#3. Second, due to the relative small size of the Steam Generator#2 leak, it was difficult to detect the effects of the leak due to the presence of the large rupture in Steam Generator#3. In fact, looking at the instruments in the nuclear control room, the pressure level was stable in Steam Generator#2 for extended periods of time. This could be interpreted as a valid data point for confirming that only Steam Generator#3 was damaged. These factors thus served as salient factors for relying only on confirmatory search.

6.3. Implications for Design

Our results suggest that the presence of and design of the procedures may have been conducive to increased risk of being influenced by these biases. In scenario 1, we found that all crews made initial efforts to obtain local samples through communications with a field operator, but only one crew abandoned these efforts after emergency procedures were entered, as the procedures did provide a follow-up. In scenario 2, several crews specifically chose not to pursue any additional actions towards diagnosing the problem, as the procedures did not require doing so. These results thus suggest that the fact that procedures did not require diagnosis of the problem, nor encouraged operators to look for alternate sources of information, dissuaded crews from exploring alternate hypotheses.

While our findings are exploratory, rather than confirmed in a deductive way, previous studies in design research have shown that interaction with written materials can cause a decrease in idea generation [18], [19], [37], which has been explicitly linked to confirmation bias in designers [38]. Furthermore, research has shown that confirmation bias influences how pilots plan decisions in adverse weather conditions [39] and how military analysts prioritise information [40]. Therefore, the design of procedures and checklists that prevent (or minimise) these biases has great potential for improving performance and, thus, safety in these fields. Research has been conducted that shows some success with regards to minimisation of design fixation [21] and debiasing decision makers through design [39], [40], as well as reducing confirmation bias in designers [38]. In particular, the use of counterfactuals, meaning examples that are opposite to the observed events or encourage thinking of the opposite of the present view, have been successful in combating confirmation bias [40], [41]. Our findings showed that successful teams employed counterfactual checks as part of their strategy, which resulted in fewer examples of biased behaviour. Therefore, we suggest that procedures created to explicitly include counterfactual checks may be successful in increasing the performance of nuclear control room operators and other operators in emergency environments through decreased confirmation bias. Another avenue for decreasing confirmation bias could be the construction of decision matrices, which have been successful in reducing confirmation bias in designers [38]. However, further research is needed to validate these proposals for operators in emergencies.

7. Conclusion

This paper investigated biased decision making in realistic extra-procedural nuclear control room scenarios. It focused on two biases that have been related to expert decision making: The first is the bias that occurs when expertise is transferred to a similar but different situation, thus causing misapplications of one's expertise. The second is confirmation bias, which is the tendency to overly prioritise and seek for information that benefits existing views. This paper presented a re-analysis of a data from two realistic simulated accidents at a pressurised water reactor, which were conducted with five nuclear control room operating crews. The scenarios required operators to perform autonomous actions outside of procedures to achieve optimal performance. We investigated whether the misapplication of expertise and confirmation bias caused suboptimal performance in the applied case of a nuclear control room emergency. While the study presented is exploratory, and the findings have thus not been validated in a deductive manner, findings in this paper provide evidence that both biases could explain differences in performance. Furthermore, we found that the use of procedures may have increased the effect of confirmation bias. We concluded by discussing implications and opportunities for the design of procedures.

Acknowledgements

We thank our research partners from the Halden HAMMLAB at the Institute for Energy Technology (IFE) in Norway for their continued support. In particular: Andreas Bye, Lars Holmgren, Salvatore Massaiu, Espen Nystad and Stine Strand. The work reported in this paper is part-funded by the OECD Halden Reactor Project.

References

[1] B. K. Burian, "Design Guidance for Emergency and Abnormal Checklists in Aviation," *Proc. Hum. Factors Ergon. Soc. 50 th Annu. Meet.*, vol. 50, no. October, pp. 1–6, 2006.

[2] B. K. Burian, "Do You Smell Smoke? Issues in the Design and Content of Checklists for Smoke, Fire, and Fumes," in *Emergency and Abnormal Situations Symposium*, 2005.

[3] N. Lau *et al.*, "Ecological interface design in the nuclear domain: An application to the secondary subsystems of a boiling water reactor plant simulator," *IEEE Trans. Nucl. Sci.*, vol. 55, no. 6, pp. 3579–3596, Dec. 2008.

[4] U. Weyer, A. O. Braseth, M. Eikås, L. Hurlen, P. Kristiansen, and J. Kvalem, "Safety Presentation in Large Screen Displays - A New Approach," in *SPE Intelligent Energy Conference and Exhibition*, 2010.

[5] A. O. Braseth and T. A. Øritsland, "Visualizing complex processes on large screen displays: Design principles based on the Information Rich Design concept," *Displays*, vol. 34, no. 3, pp. 215–222, 2013.

[6] J. Wreathall and J. Reason, "Latent Failures and Human Performance in Significant Operating Events," *JWCo. Inc*, 1993.

[7] U.S. Nuclear Regulatory Commission, "Review of Findings for Human Performance Contribution to Risk in Operating Events (NUREG/CR-6753, INEEL/EXT-01-01166)," 2002.

[8] S. Massaiu, “Critical features of emergency procedures: Empirical insights form simulations of nuclear power plant operation,” *Training*, pp. 277–284, 2010.

[9] P. Du and E. F. MacDonald, “Products’ Shared Visual Features Do Not Cancel in Consumer Decisions,” *J. Mech. Des.*, vol. 137, no. 7, p. 71409, 2015.

[10] P. Du and E. F. MacDonald, “Eye-Tracking Data Predicts Importance of Product Features and Saliency of Size Change,” *Vol. 5 25th Int. Conf. Des. Theory Methodol. ASME 2013 Power Transm. Gearing Conf.*, vol. 136, no. 8, p. V005T06A024, 2013.

[11] K. Goucher-Lambert and J. Cagan, “The Impact of Sustainability on Consumer Preference Judgments (DETC2014-34739),” *ASME 2014 Int. Des. Eng. Tech. Conf. Comput. Inf. Eng. Conf.*, vol. 137, no. c, pp. 1–11, Aug. 2014.

[12] K. Goucher-Lambert, J. Moss, and J. Cagan, “Inside the Mind: Using Neuroimaging to Understand Moral Product Preference Judgments Involving Sustainability,” *J. Mech. Des.*, vol. 139, no. 4, p. 41103, 2017.

[13] S. J. Faerber and C. C. Carbon, “Jump on the innovator’s train: Cognitive principles for creating appreciation in innovative product designs,” *Res. Eng. Des.*, vol. 24, no. 3, pp. 313–319, 2013.

[14] T. N. Reid, E. F. MacDonald, and P. Du, “Impact of Product Design Representation on Customer Judgment,” *J. Mech. Des.*, vol. 135, no. 9, p. 91008, 2013.

[15] S. T. Wei, L. C. Ou, M. R. Luo, and J. B. Hutchings, “Package design: Colour harmony and consumer expectations,” *Int. J. Des.*, vol. 8, no. 1, pp. 109–126, 2014.

[16] M. Perez Mata, S. Ahmed-Kristensen, P. B. Brockhoff, and H. Yanagisawa, “Investigating the influence of product perception and geometric features,” *Res. Eng. Des.*, vol. 28, no. 3, pp. 357–379, 2017.

[17] E. Andersen and A. Maier, “The attentional capture of colour in visual interface design: A controlled-environment study,” in *Proceedings of the International Conference on Engineering Design, ICED*, 2017, vol. 8, no. DS87-8.

[18] D. G. Jansson and S. M. Smith, “Design fixation,” *Des. Stud.*, vol. 12, no. 1, pp. 3–11, 1991.

[19] L. A. Vasconcelos and N. Crilly, “Inspiration and fixation: Questions, methods, findings, and challenges,” *Des. Stud.*, vol. 42, pp. 1–32, 2016.

[20] J. B. Nikander, L. A. Liikkanen, and M. Laakso, “The preference effect in design concept evaluation,” *Des. Stud.*, vol. 35, no. 5, pp. 473–499, Sep. 2014.

[21] V. K. Viswanathan and J. S. Linsey, “Design Fixation and Its Mitigation: A Study on the Role of Expertise,” *J. Mech. Des.*, vol. 135, no. 5, p. 51008, Apr. 2013.

[22] D. Kahneman and G. Klein, “Conditions for intuitive expertise: A failure to disagree.,” *Am. Psychol.*, vol. 64, no. 6, pp. 515–526, 2009.

[23] G. Gigerenzer and W. Gaissmaier, “Heuristic Decision Making,” *Annu.*

Rev. Psychol., vol. 62, no. 1, pp. 451–482, 2011.

[24] S. Massaiu and L. Holmgren, “Diagnosis and Decision-Making with Emergency Operating Procedures in Non-typical Conditions: A HAMMLAB Study with U.S. Operators,” *HWR-1121. Halden, Norw. OECD Halden React. Proj.*, 2014.

[25] S. Massaiu and L. Holmgren, “A HAMMLAB HRA Data Collection with U.S. Operators. HWR-1123. Halden, Norway: OECD Halden Reactor Project.,” *HWR-1123. Halden, Norw. OECD Halden React. Proj.*, 2016.

[26] R. S. Nickerson, “Confirmation bias: A ubiquitous phenomenon in many guises.,” *Rev. Gen. Psychol.*, vol. 2, no. 2, pp. 175–220, 1998.

[27] H. A. Simon, “What Is an Explanation of Behavior?,” *Psychol. Sci.*, vol. 3, no. 3, pp. 150–161, 1992.

[28] J. Shanteau, “Competence in experts: The role of task characteristics,” *Organ. Behav. Hum. Decis. Process.*, vol. 53, no. 2, pp. 252–266, Nov. 1992.

[29] J. Shanteau, “How much information does an expert use? Is it relevant?,” *Acta Psychol. (Amst).*, vol. 81, no. 1, pp. 75–86, 1992.

[30] S. E. Dreyfus, “The Five-Stage Model of Adult Skill Acquisition,” *Bull. Sci. Technol. Soc.*, vol. 24, no. 3, pp. 177–181, 2004.

[31] J. S. Bruner and M. C. Potter, “Interference in Visual Recognition,” *Science (80-.).*, vol. 144, no. 3617, pp. 424–425, Apr. 1964.

[32] L. Ross and C. A. Anderson, “Shortcomings in the attribution process: On the origins and maintenance of erroneous social assessments,” *Judgm. under Uncertain. Heuristics biases*, no. 1977, pp. 129–153, 1982.

[33] C. a. Anderson, M. R. Lepper, and L. Ross, “Perseverance of social theories: The role of explanation in the persistence of discredited information.,” *J. Pers. Soc. Psychol.*, vol. 39, no. 6, pp. 1037–1049, 1980.

[34] J. Baron, “Myside Bias in Thinking About Abortion,” *Think. Reason.*, vol. 1, no. 3, pp. 221–235, 1995.

[35] P. C. Wason, “On the failure to eliminate hypotheses in a conceptual task,” *Q. J. Exp. Psychol.*, vol. 12, no. 3, pp. 129–140, 1960.

[36] D. Griffin and A. Tversky, “The weighing of evidence and the determinants of confidence,” *Cogn. Psychol.*, vol. 24, no. 3, pp. 411–435, 1992.

[37] L. A. Vasconcelos, C. C. Cardoso, M. Sääksjärvi, C.-C. Chen, and N. Crilly, “Inspiration and Fixation: The Influences of Example Designs and System Properties in Idea Generation,” *J. Mech. Des.*, vol. 139, no. 3, p. 31101, 2017.

[38] G. M. Hallihan, H. Cheong, and L. H. Shu, “Confirmation and Cognitive Bias in Design Cognition,” in *Volume 7: 9th International Conference on Design Education; 24th International Conference on Design Theory and Methodology*, 2012, p. 913.

[39] S. Walmsley and A. Gilbey, “Debiasing visual pilots’ weather-related decision making,” *Appl. Ergon.*, vol. 65, pp. 200–208, 2017.

[40] M. B. Cook and H. S. Smallman, “Human Factors of the Confirmation Bias in Intelligence Analysis: Decision Support From Graphical Evidence

Landscapes," *Hum. Factors J. Hum. Factors Ergon. Soc.*, vol. 50, no. 5, pp. 745–754, 2008.

[41] A. D. Galinsky and G. B. Moskowitz, "Counterfactuals as Behavioral Primes: Priming the Simulation Heuristic and Consideration of Alternatives," *J. Exp. Soc. Psychol.*, vol. 36, no. 4, pp. 384–409, 2000.

CHAPTER 5:

Conclusion

5.1. Chapter Introduction

The research presented in this PhD thesis represented an effort in using cognitive theory to improve the design of visual interfaces for use in emergencies. This was reflected in both original research and reviews, which aimed to answer the research question presented in section 1.3.:

- How can visual design for emergencies be improved through improving design research [1], and through cognitive theory on attention [2], biases and heuristics [3]?

The answer to the research question was explored through three angles, that are reflected in the research question: 1) Improving design research, 2) applying attention theory to design, and 3) applying bias and heuristics literature to design.

In this chapter, the contributions of this thesis for each angle will be summarised, and suggestions for possible future work will be made based on the findings.

5.2. Improving Design Research

The first angle explored was to improve design through improving design research. This angle was explored through three approaches: reviewing the design literature and offering concrete suggestions for improvement (Paper 1/ Andersen et al., submitted), reviewing cognitive theory and suggesting possible impacts for design research (all papers), and creating new experimental paradigms that address methodological issues with previous paradigms (Papers 3, 4 and 5/ Andersen & Maier, 2017; submitted; Andersen et al., submitted).

For example, Paper 5 (Andersen et al., submitted) employed a fully automated experimental design, featuring variations of a visual display that was based on both a realistic practice case and the displays used in experimental psychology. The adopted approach allowed for a close relation between theory building and application, which may eventually enable design research to establish strong theories for the visual features. The fully automated nature of the experiment further allows for easier replication of the experiment, which could both increase validity of the results, and offer easily comparable results for a number of cultural or professional groups, thereby also increasing ecological validity and possibility for improved theory building (Cash, 2018).

As argued in Paper 1 (Andersen et al., submitted), future studies are needed that a) reproduces and/or builds on previous studies through either elaborating the topic or providing rigorously validated methods and measures for the constructs, and b) develop a common theoretical basis for the constructs to allow for more systematic exploration and testing. As was further argued in Paper 1 (Andersen et al., submitted), and as all research in this thesis serves an example of, achieving these goals could be eased by adopting and adapting pre-existing cognitive theory and methodology.

5.3. Applying Attention Theory to Design

The second angle explored was to apply attention theory to design. This was done through reviewing attention theory and offering suggestions on how to apply the findings in design research and design practice, and through original experimental research that tested specific predictions from attention theory in more complex displays that better resembles the interfaces used by nuclear control room operators and other professionals in emergencies.

The review of attention theory (Paper 2/ Andersen & Maier, 2016) outlined how attention theory can be used to predict attentional allocation based on the factors that capture and guide gaze, and the factors that modulate this effect. The presented experimental papers (Papers 3, 4 and 5/ Andersen & Maier, 2017; submitted; Andersen et al., submitted) tested many of these factors in complex visual interfaces that closer resemble the interfaces used by operators and other professionals in emergencies. Specifically, the effect on attention of specific colours (Papers 3, 4, and 5/ Andersen & Maier, 2017; submitted; Andersen et al., submitted), cognitive load, visual load, visual dilution and distractors (Paper 5/ Andersen et al., submitted) on attentional allocation were studied.

Taken together, these findings can inform designers about how attention will be located depending on how they use these factors in their interfaces, and depending on the environment that they predict the interfaces to be used in. For example, the findings can be applied to align the importance of visual artefacts and their colour according to the colour's attentional capture, which was shown to be particularly important when designing complex interfaces that cause high visual load.

However, as seen in Paper 5 (Andersen et al., submitted), all theory does not always translate to reality. For example, we found that the Load Theory of attention was limited in its generalizability from the simplified interfaces of experimental psychology to the more complex interface in our research. This indicates that there is high potential value for both attention theory and design in testing attention theory in more complex and/or realistic interfaces. For attention theory, it could provide ecological validity tests of the experimental findings. For design practice and research, it could lead to a better understanding of how interfaces will be used, and therefore improved visual designs. While an extensive endeavour, the accumulation of findings from similar use cases may someday lead to the establishment of generalisable principles for visual design that additionally further our understanding of the underlying cognitive mechanisms. Furthermore, making designs for real-life use requires considering all the levels of use, as pointed out by Rasmussen (1983). Therefore, future studies are needed that link in more detail the association between biases at the attention level (which, using Rasmussen's terminology, falls at the skills level of performance) and biases at the decision-making level (which, using Rasmussen's terminology, usually falls at the rules or knowledge levels of performance).

5.4. Applying Bias and Heuristics Literature to Design

The third angle explored was to apply bias and heuristic literature to design. This was done through selecting two biases that have commonly been found to influence expert decision making and investigating a) whether they influenced decision making of operators of nuclear control room operators, and b) whether the visual procedures that operators rely on amplified these effects.

These findings were presented in Paper 6 (Andersen et al., 2018), which presented an analysis of previously collected data on the performance of five crews in two simulated emergencies. The behavioural patterns of higher, medium, and lower performing crews were analysed and it was found that lower performing crews may have in part received a lower score due to being affected by confirmation bias and/or bias of misapplied expertise. Furthermore, it was found that these effects may have been amplified by the design of the procedures that operators rely on for the decision making.

These findings can thus be translated into potential improvements for how the procedures could be designed to prevent these biases and improve visual search through design. For example, the results showed that implementing counterfactuals in procedure design, or prompts for considering the opposite hypothesis, could be particularly effective due to their effectiveness in preventing confirmation bias in other domains (Cook & Smallman, 2008; Galinsky & Moskowitz, 2000).

The study presented in this thesis is thus an example of how application of bias and heuristics literature to a design relevant case can provide a greater understanding of how interfaces are used, which in turn can be used to improve them. As the study was exploratory, however, future studies need to be conducted that test and confirm the findings. Furthermore, future studies could test whether the proposed interventions for combating confirmation bias could also be viable for concrete use cases such as nuclear control room operator decision making. Finally, while the biases were selected due to their relevance to expert decision making in emergencies, many other biases and heuristics have been established in the literature that could have similar applications for improving visual design for emergencies. In addition to the aforementioned need for studies on the association between attention bias and decision-making bias, future studies are thus needed to give a better understanding on the interaction between biased behaviour and visual guidelines.

5.5. Conclusion

In sum, this thesis has proposed that visual design for emergencies can be improved through 1) improving design research by applying pre-existing cognitive theory and methodology, 2) applying and testing attention theory in complex visual interfaces that better resemble the real displays used in emergencies, and 3) applying and testing the effects of heuristics and biases on decision making of operators in simulation studies and investigating whether the design of the employed visual guidelines amplify these effects. While more research is needed, as always, it is my hope that the findings in this thesis may contribute to making better visual interfaces, and thus improving visual information search in emergencies.

References

Andersen, E., Kozin, I., & Maier, A. (2018). Biased Decision Making in Realistic Extra-Procedural Nuclear Control Room Scenarios. In *8th International Conference on Design Computing and Cognition*.

Andersen, E., & Maier, A. (2016). What captures gaze in visual design? Insights from cognitive psychology. In *Proceedings of NordDesign, NordDesign 2016* (Vol. 1).

Andersen, E., & Maier, A. (2017). The attentional capture of colour in visual interface design: A controlled-environment study. In *Proceedings of the International Conference on Engineering Design, ICED* (Vol. 8).

Baars, B. J., & Franklin, S. (2007). An architectural model of conscious and unconscious brain functions: Global Workspace Theory and IDA. *Neural Networks*, *20*(9), 955–961. https://doi.org/10.1016/j.neunet.2007.09.013

Benoni, H., & Tsal, Y. (2010). Where have we gone wrong? Perceptual load does not affect selective attention. *Vision Research*, *50*(13), 1292–1298. https://doi.org/10.1016/j.visres.2010.04.018

Biggs, A. T., Kreager, R. D., & Davoli, C. C. (2015). Finding a link between guided search and perceptual load theory. *Journal of Cognitive Psychology*, *27*(2), 164–179. https://doi.org/10.1080/20445911.2014.987676

Blessing, L. T. M., & Chakrabarti, A. (2009). *DRM: A Design Reseach Methodology*. *DRM, a Design Research Methodology*. Springer. https://doi.org/10.1007/978-1-84882-587-1_2

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, *97*(4), 523–547. https://doi.org/10.1037/0033-295X.97.4.523

Bundesen, C., & Habekost, T. (2008). *Principles of Visual Attention: Linking Mind and Brain* (lt;oxdc:e). Oxford University Press.

Burian, B. K. (2005). Do You Smell Smoke? Issues in the Design and Content of Checklists for Smoke, Fire, and Fumes. In *Emergency and Abnormal Situations Symposium*. Retrieved from http://humansystems.arc.nasa.gov/flightcognition/Publications/Burian_ISASI_2005_paper.pdf

Cash, P. J. (2018). Developing theory-driven design research. *Design Studies*, *56*, 84–119. https://doi.org/10.1016/j.destud.2018.03.002

Cook, M. B., & Smallman, H. S. (2008). Human Factors of the Confirmation Bias in Intelligence Analysis: Decision Support From Graphical Evidence Landscapes. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(5), 745–754. https://doi.org/10.1518/001872008X354183

Dehaene, S., Charles, L., King, J. R., & Marti, S. (2014). Toward a computational theory of conscious processing. *Current Opinion in Neurobiology*, *25*, 76–84. https://doi.org/10.1016/j.conb.2013.12.005

Dinar, M., Shah, J. J., Cagan, J., Leifer, L., Linsey, J., Smith, S. M., & Hernandez, N. V. (2015). Empirical Studies of Designer Thinking: Past, Present, and Future. *Journal of Mechanical Design*, *137*(2), 21101. https://doi.org/10.1115/1.4029025

Dreyfus, S. E. (2004). The Five-Stage Model of Adult Skill Acquisition. *Bulletin of Science, Technology & Society*, *24*(3), 177–181. https://doi.org/10.1177/0270467604264992

Evans, J. S. B. T. (2003). In two minds: dual-process accounts of reasoning. *Trends in Cognitive Sciences*, *7*(10), 454–459. https://doi.org/10.1016/j.tics.2003.08.012

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annu. Rev. Psychol.*, *59*, 255–278.

Faerber, S. J., & Carbon, C. C. (2013). Jump on the innovator's train: Cognitive principles for creating appreciation in innovative product designs. *Research in Engineering Design*, *24*(3), 313–319. https://doi.org/10.1007/s00163-012-0148-7

Fernandes, A., & Braarud, P. Ø. (2015). Exploring Measures of Workload, Situation Awareness, and Task Performance in the Main Control Room. *Procedia Manufacturing*, *3*, 1281–1288. https://doi.org/10.1016/j.promfg.2015.07.273

Gailliot, M. T., Baumeister, R. F., DeWall, C. N., Maner, J. K., Plant, E. A., Tice, D. M., … Schmeichel, B. J. (2007). Self-control relies on glucose as a limited energy source: Willpower is more than a metaphor. *Journal of Personality and Social Psychology*, *92*(2), 325–336. https://doi.org/10.1037/0022-3514.92.2.325

Galinsky, A. D., & Moskowitz, G. B. (2000). Counterfactuals as Behavioral Primes: Priming the Simulation Heuristic and Consideration of Alternatives. *Journal of Experimental Social Psychology*, *36*(4), 384–409. https://doi.org/10.1006/jesp.1999.1409

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic Decision Making. *Annual Review of Psychology*, *62*(1), 451–482. https://doi.org/10.1146/annurev-psych-120709-145346

Hergenhahn, B. R., & Henley, T. B. (2014). Hergenhahn's An Introduction to the History of Psychology, 696.

Kahneman, D. (2011). *Thinking, fast and slow* (Vol. 1). Farrar, Straus and Giroux New York.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. *Heuristics and Biases: The Psychology of Intuitive Judgment*, *49*, 81.

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist*, *64*(6), 515–526. https://doi.org/10.1037/a0016755

Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology: General*, *139*(4), 665–682. https://doi.org/10.1037/a0020198

Kyllingsbæk, S., Sy, J. L., & Giesbrecht, B. (2011). Understanding the allocation of attention when faced with varying perceptual load in partial report: A computational approach. *Neuropsychologia*, *49*(6), 1487–1497. https://doi.org/10.1016/j.neuropsychologia.2010.11.039

Lamme, V. A. F. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences*, *7*(1), 12–18. https://doi.org/10.1016/S1364-6613(02)00013-X

Lavie, N., Hirst, A., De Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*(3), 339–354. https://doi.org/10.1037/0096-3445.133.3.339

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics*, *56*(2), 183–197.

https://doi.org/10.3758/BF03213897

Muraven, M., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle? *Psychological Bulletin*, *126*(2), 247.

Muraven, M., Tice, D. M., & Baumeister, R. F. (1998). Self-control as a limited resource: Regulatory depletion patterns. *Journal of Personality and Social Psychology*, *74*(3), 774.

Nørager, R. (2009). *Low Level Cognition in User Interfaces*. Department of Psychology, University of Aarhus.

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion Drives Attention: Detecting the Snake in the Grass. *Journal of Experiemntal Psychology: General*, *130*(3), 466–478. https://doi.org/10.1037/AXJ96-3445.130.3.466

Papalambros, P. Y. (2015). Design Science: Why, What and How. *Design Science*, *1*, e1. https://doi.org/10.1017/dsj.2015.1

Popper, E. R. (1959). *The Logic of Scientific Discovery*. *Physics Today* (Vol. 12). Routledge. https://doi.org/10.1063/1.3060577

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics*, *SMC-13*(3), 257–266. https://doi.org/10.1109/TSMC.1983.6313160

Shanteau, J. (1992). Competence in experts: The role of task characteristics. *Organizational Behavior and Human Decision Processes*, *53*(2), 252–266. https://doi.org/10.1016/0749-5978(92)90064-E

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, *74*(11), 1–29. https://doi.org/10.1037/h0093759

Sternberg, S. (1966). High-Speed Scanning in Human Memory. *Science*, *153*(3736), 652–654.

Theeuwes, J. (2010). Top–down and bottom–up control of visual selection. *Acta Psychologica*, *135*(2), 77–99. https://doi.org/10.1016/j.actpsy.2010.02.006

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive*

Psychology, *12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Tsal, Y., & Benoni, H. (2010). Diluting the Burden of Load: Perceptual Load Effects Are Simply Dilution Effects. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6), 1645–1656. https://doi.org/10.1037/a0018172

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131. Retrieved from http://psiexp.ss.uci.edu/research/teaching/Tversky_Kahneman_1974.pdf

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453–458. https://doi.org/10.1126/science.7455683

U.S. Nuclear Regulatory Commission. (2002). *Review of Findings for Human Performance Contribution to Risk in Operating Events (NUREG/CR-6753, INEEL/EXT-01-01166).* Retrieved from http://www.nrc.gov/reading-rm/doc-collections/nuregs/contract/cr6753/

Vasconcelos, L. A., & Crilly, N. (2016). Inspiration and fixation: Questions, methods, findings, and challenges. *Design Studies*, *42*, 1–32. https://doi.org/10.1016/j.destud.2015.11.001

Walmsley, S., & Gilbey, A. (2017). Debiasing visual pilots' weather-related decision making. *Applied Ergonomics*, *65*, 200–208. https://doi.org/10.1016/j.apergo.2017.06.019

Wolfe, J. M. (2007). Guided search 4.0. *Integrated Models of Cognitive Systems*, (3), 99–120. https://doi.org/10.1007/978-94-011-5698-1_30

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 419–433.

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, *5*(6), 495–501. https://doi.org/10.1038/nrn1411

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*(3), 58. https://doi.org/10.1038/s41562-017-0058