Depicting Movement Data with Animations for Embodied and Real-Time Decision-Making

A User Study with Air Traffic Control Displays and Real-Time Movement Data

---

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Sara Mary Rosa Maggi

von Locarno TI

Promotionskommission

Prof. Dr. Sara Irina Fabrikant (Vorsitz)
Prof. Dr. Robert Weibel
Prof. Dr. Thomas Shipley
Prof. Dr. Christophe Hurter

Zürich, 2017

ABSTRACT

How to effectively and efficiently represent the dynamics of spatial phenomena and processes has been a long-standing research question in geographic information science (GIScience). In a digital information age, computer-generated animations that depict movement data have become increasingly popular, as they apparently visualize real-world spatio-temporal movement changes with corresponding changes over time in a moving display. Animation thus seems to be a suitable display method for facilitating the recognition of spatio-temporal movement patterns and the prediction of future spatio-temporal events.

However, the manner by which animations are designed may limit the effectiveness and efficiency of visuospatial decision-making. Furthermore, the specific decision-making task or context of use, as well as the viewer's perceptual, cognitive and affective background might also influence visuospatial decision-making with animations. These factors are not well understood to date. More empirical studies, as well as new methods to evaluate animations, are thus needed.

This work proposes a user-centred empirical approach to evaluate animation design characteristics for space-time decision-making with movement data. Two experiments are conducted with the overall aim of answering the following main research question: How should animations of real-time movement data be designed considering the task and/or use contexts, and user characteristics? More specifically, we test the influence of the three main visual analytics (VA) dimensions on viewer spatio-temporal decision-making with animations: (1) the use context and respective task characteristics, (2) the animation display design, and (3) user characteristics. To test each respective dimension, we undertook the following investigations:

(1) Using current air traffic control (ATC) scenarios and existing ATC displays we empirically investigated how aircraft movement changes and future aircraft movement patterns can be visualized for effective and efficient decision-making in ATC.

(2) We empirically investigated how movement characteristics (i.e., acceleration, heading direction, etc.) can be depicted, and how animation design (i.e., continuous vs. semi-static animations) might influence viewer task performances.

(3) We empirically investigated how perceptual, cognitive, and affective characteristics of viewers (i.e., expertise, spatial abilities, stress or motivation) might influence visuospatial decision-making with animations.

We approached these questions through novel empirical data triangulation that integrates psychophysical sensing (i.e., electrodermal responses (EDA)), brain activity (i.e., electroencephalography (EEG)), and eye tracking (ET) with standardized questionnaires.

The results of the experiments showed that these three factors (i.e., the use context and respective task characteristics, the animation display design, and the user characteristics) indeed influence visuospatial decision-making using animations of aircraft movement data. We found that viewer decision-making was affected by animation design depending on expertise and task type. Unsurprisingly, ATC experts performed typical ATC tasks more accurately compared to novices. However, the task performance of the experts differed between continuous animation and semi-static animation designs depending on the ATC task. Surprisingly, experts responded more accurately with the novel continuous animation designs compared to the semi-static animations that are more familiar to them in critical ATC tasks for predicting future aircraft movements. In apprehension tasks of aircraft movement changes, experts performed in similar ways with both animation designs. Moreover, viewer characteristics, such as spatial abilities and emotional aspects including engagement or motivation, seemed to affect viewer task performances as well. Higher-spatial and more engaged (or more motivated) viewers performed both tasks more effectively than lower-spatial decision makers and less-engaged (or less-motivated) viewers.

Overall, our unique empirical results related to the depiction of real-time movement data contribute to GIScience and cartography in two important ways. First, we are beginning to better understand how viewer mental processes, including perception and cognition, as well as their affective states might influence the effectiveness and efficiency of visuospatial decision-making with animations. Second, we are now able to derive empirically validated design guidelines for perceptually salient, affectively engaging, and cognitively inspired animations.

ZUSAMMENFASSUNG

Die Frage, wie die Dynamik räumlicher Phänomene und Prozesse effektiv und effizient dargestellt werden kann, ist seit vielen Jahren ein Forschungsthema in der Geographischen Informationswissenschaft (GIScience). Im digitalen Informationszeitalter erfreuen sich computergenerierte Animationen für die Darstellung von Bewegungsdaten zunehmender Beliebtheit, da sie augenscheinlich reale räumlich-zeitliche Bewegungsänderungen mit den entsprechenden zeitlichen Veränderungen in einer bewegten Anzeige visualisieren. Animationen scheinen also eine geeignete Darstellungsmethode zu sein, um die Erkennung räumlich-zeitlicher Bewegungsmuster und die Vorhersage künftiger räumlich-zeitlicher Ereignisse zu vereinfachen.

Die Art und Weise der Animationsgestaltung könnte jedoch die Wirksamkeit und Effizienz der visuell-räumlichen Entscheidungsfindung einschränken. Darüber hinaus könnten die spezifische Entscheidungsaufgabe oder der Anwendungskontext und der perzeptive, kognitive und affektive Hintergrund des Betrachters die visuell-räumliche Entscheidungsfindung auf Grundlage von Animationen beeinflussen. Diese Faktoren werden bislang nur unzureichend verstanden. Es sind daher weitere empirische Studien sowie neue Methoden zur Bewertung von Animationen erforderlich.

In dieser Arbeit wird ein anwenderorientierter empirischer Ansatz zur Beurteilung der Animationsgestaltungsmerkmale für die Raum-Zeit-Entscheidungsfindung mit Bewegungsdaten vorgeschlagen. Es werden zwei Experimente mit dem Gesamtziel durchgeführt, die folgende zentrale Forschungsfrage zu beantworten: Wie sollen Animationen von Echtzeitbewegungsdaten unter Berücksichtigung des Aufgaben- und / oder Anwendungskontexts und der Anwendermerkmale gestaltet werden? Im Einzelnen testen wir den Einfluss der drei VA-Hauptdimensionen (Visual Analytics) auf die räumlich-zeitliche Entscheidungsfindung des Betrachters mit Animationen: (1) den Anwendungskontext und die jeweiligen Aufgabenmerkmale, (2) das Animationsdisplay-design und (3) die Anwendermerkmale.

Mithilfe von aktuellen Flugsicherungsszenarien (FSS) und existierenden Flugsicherungsanzeigen untersuchen wir empirisch, wie sich Flugzeugbewegungsänderungen und zukünftige Flugzeugbewegungsmuster für eine effektive und effiziente Entscheidungsfindung bei der Flugsicherung visualisieren lassen.

Wir untersuchen empirisch, wie sich Bewegungsmerkmale (d. h. Beschleunigung, Kursrichtung usw.) darstellen lassen und wie sich die Animationsgestaltung (d. h. kontinuierliche vs. semistatische Animationen) auf die Aufgabenausführung des Betrachters auswirken könnte. Wir untersuchen empirisch, wie sich die perzeptiven, kognitiven und affektiven Merkmale des Betrachters (d. h. Sachverstand, räumliche Fähigkeiten, Stress oder Motivation) auf die visuell-räumliche Entscheidungsfindung mit Animationen auswirken können. Wir nähern uns diesen Fragen durch eine neue empirische Datentriangulationsmethode an, welche die psychophysische Wahrnehmung (d. h. elektrodermale Reaktionen, EDA), die Hirnaktivität (d. h. Elektroenzephalographie, EEG) und Augenverfolgung (Eyetracking, ET) mit standardisierten Fragebögen integriert.

Die Ergebnisse der Experimente zeigen, dass diese drei Faktoren (d. h. der Anwendungskontext und die jeweiligen Aufgabeneigenschaften, das Animationsdisplaydesign und die Anwendermerkmale) tatsächlich die visuell-räumliche Entscheidungsfindung mit Animationen bei Flugzeugbewegungsdaten beeinflussen. Wir haben festgestellt, dass die Entscheidungsfindung des Betrachters je nach Fachwissen und Aufgabenart durch das Animationsdesign beeinflusst wird. Kaum überraschend führten Flugsicherungsexperten die typischen Flugsicherungsaufgaben im Vergleich zu Anfängern mit einer höheren Genauigkeit durch. Die Leistungsfähigkeit der Experten weicht jedoch je nach Flugsicherungsaufgabe zwischen kontinuierlicher Animation und halbstatischen Animationsdesigns voneinander ab. Überraschenderweise reagieren Experten mit den neuartigen kontinuierlichen Animationsdesigns im Vergleich zu den semistatischen Animationen, die ihnen bei kritischen Flugsicherungsaufgaben für die Vorhersage zukünftiger Flugzeugbewegungen mehr vertraut sind, mit einer höheren Genauigkeit. Im Hinblick auf Auffassungsaufgaben bezüglich Flugzeugbewegungsänderungen zeigen Experten bei beiden Animationsdesigns ein ähnliches Verhalten. Darüber hinaus scheinen sich Betrachtermerkmale wie etwa räumliche Fähigkeiten und emotionale Aspekte, einschließlich von Engagement oder Motivation, auf die Leistungsfähigkeit von Betrachtern auszuwirken. Engagiertere (oder motiviertere) Betrachter mit einem besseren räumlichen Denkvermögen erfüllten beide Aufgaben effektiver als Entscheidungsträger mit einem geringeren räumlichen Denkvermögen und weniger engagierte (oder weniger motivierte) Betrachter.

Insgesamt leisten unsere einzigartigen empirischen Ergebnisse im Zusammenhang mit der Darstellung von Echtzeitbewegungsdaten in zweierlei Hinsicht einen Beitrag für die

GIScience und für die Kartografie. Zum einen beginnen wir, zu verstehen, wie sich die mentalen Prozesse bei Betrachtern, einschließlich von Wahrnehmung und Kognition sowie der affektiven Zustände, auf die Effektivität und Effizienz visuell-räumlicher Entscheidungen mit Animationen auswirken können. Zum anderen sind wir nun dazu in der Lage, empirisch validierte Gestaltungsrichtlinien für perzeptiv markante, affektiv ansprechende und kognitiv inspirierte Animationen abzuleiten.

ACKNOWLEDGEMENTS

During these four years of my PhD, I met different people with whom I had the opportunity to collaborate and who supported me in different ways to complete the work presented in this thesis. I would like to express my special thanks to the following people:

First of all, I would like to give my deep appreciation and thanks to my supervisor, Prof. Dr. Sara Irina Fabrikant, for providing me the remarkable opportunity to take part in the AniMOVE research project and to join the GIVA group of the University of Zurich, for her essential support during the entire PhD process, and for her always prompt availability to discuss and advise research ideas and teaching topics with me. Your invaluable advice has been very important for me in my growth as a researcher and teacher, as well as for my future career.

Second, I would also like to thank all of my committee members, Prof. Dr. Robert Weibel (University of Zurich, Switzerland), Prof. Dr. Thomas Shipley (Temple University, USA), and Prof. Dr. Christophe Hurter (ENAC, France), for your invaluable collaboration in the design and realisation of this research project, especially for the two user studies, and for your important suggestions and comments.

Additionally, I would like to thank all of my colleagues of the GIVA group, composed by Dr. Arzu Çöltekin, Dr. Tumasch Reichenbacher, Dr. Kai-Florian Richter, Dr. Alžběta Brychtová, Dr. Jan Wilkening, Kenan Bektas, Gianluca Boo, Annina Brügger, André Bruggmann, Sascha Credé and Ismini Lokka for your support and discussions. I particularly thank André Bruggmann, my office mate, for his helpful advice and for always motivating me with his good mood. I would like to extend a big thanks to Gianluca, Ismini, and Ekaterina Egorova. You always had an open ear for me, and you motivated me a lot during my PhD. Thanks also for the nice moments spent together outside of working time.

From ENAC (France), I would especially like to thank Prof. Dr. Christophe Hurter and Dr. Jean-Paul Imbert for your invaluable help in designing and realising the two experiments with air traffic controllers. A special thanks goes to Dr. Maxime Cordail, as well, for your precious time and for helping me significantly with programming simulated air traffic animated displays. From Masaryk University, I would like to express my appreciations especially to Dr. Zbyněk Štěrba and Dr. Čeněk Šašinka for sharing your time and your expertise in user studies, cognitive science, and cartography. Thank you to all of

you, and to Prof. Dr. Shipley and Kelly Bower, as well, for assisting me extensively in the organisation of the two experiments of this thesis at ENAC (Toulouse, France), Temple University (Philadelphia, USA), and Brno Airport (Czech Republic); for helping me in finding expert participants; and for providing the adequate structure to run the experiments.

I would like to thank Dr. Benny Briesemeister, Prof. Dr. Andrew Duchowski, Prof. Dr. Krzysztof Krejtz, Dr. Georgios Papastefanou, and Bruno Rütsche for your expertise in EEG, eye movement, and EDA analysis, which provided an essential foundation to the data analysis component of my work. Additionally, I would like to extend my thanks to my interview partners from ENAC (Toulouse, France), ETH Zurich, and in the floorball domain for your important availability in discussing the main topics of my thesis, for your expertise, and to the EHT Zurich for the permission to use its soccer data for preparing the concept of my thesis.

I would like to thank all of the air traffic controllers from Skyguide (Zurich), ENAC (Toulouse, France), and Brno (Czech Republic), as well as all of the students and collaborators from ENAC (Toulouse, France), Temple University (Philadelphia, USA), and Zurich University who participated in the pilot and main user studies of this thesis.

A special and huge thanks goes to my life-partner, my family, and all my friends for immensely supporting me during the whole PhD process, your encouragement, and your love. I am very grateful to my mother, Manuela, my father, Marco, and my brother, Stefano, for your inspiration, confidence, and motivation. I would also like to thank all of my friends (especially Veronica, Miriam and Anna) who, with your smiles and interesting activities outside of work time, allowed me to complete my PhD with a minimum of effort and worries. Last but not least, I would like to thank my life-partner, Michael, with whom I have grown up and have spent more than one-third of my life, and who, including his family, Françoise, Alexandra and Bruno, always supports me with an abundance of patience, love, and new inspirations.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE

INTRODUCTION

In many research and application fields, such as in the emergency management and surveillance domains, animations of spatio-temporal events are an effective and efficient means for supporting users in their spatio-temporal decisions or for data exploration (Schlienger et al. 2007; Kobayashi et al. 2011; Calderon et al. 2014). Animations with appropriate visual and dynamic variables might not only facilitate users in creating a better mental representation of the current dynamics of the depicted spatio-temporal phenomena, but they might also help users in anticipating future developments of spatio-temporal events. According to the *Congruence Principle* for creating effective graphics (Tversky et al. 2002), animations consistently and realistically represent the movement dynamics of moving objects, such as the movement of people, animals, or ideas.

Movement dynamics describes how moving objects translate through space over time. As Peuquet (2002) indicated, the movement and spatio-temporal patterns of moving objects can be reduced into three main components: space (where), time (when), and objects (what), as shown in Figure 1. In addition, the movement of a certain moving object is not isolated but is embedded into a specific geographical context that might interact with, influence, cause and explain its specific behaviour and movement dynamics (i.e., by external factors or by other moving objects). Moving objects can be categorized according to *generic spatio-temporal patterns* (e.g., change in direction or in speed) and to *behavioural movement patterns* related to a specific geographical context (e.g., animals that are fighting or playing) (Dodge et al. 2008).

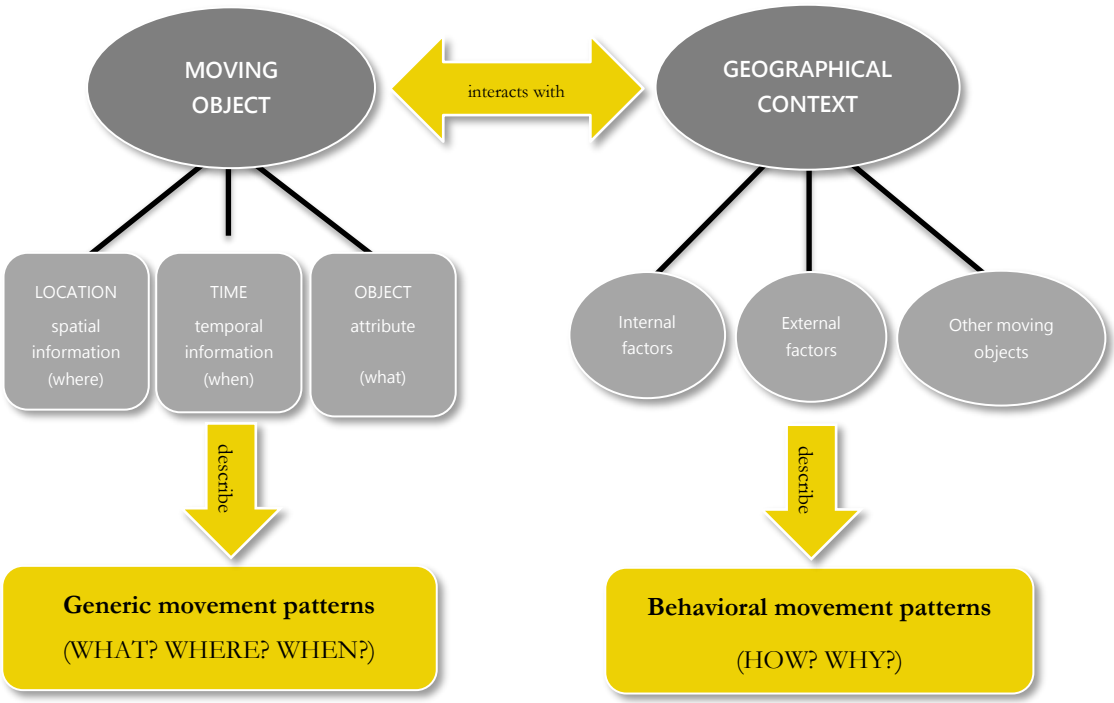


Figure 1: *Characteristics of a movement object: location, time, and object and its geographic context (after Andrienko et al. 2011; Dodge et al. 2008).*

In GIScience, the analysis and representation of movement patterns is a relatively young subfield that has until now mostly focused on the computational extraction of relevant and meaningful information from these data (Andrienko et al. 2007b). It was developed from static cartography (Hägerstrand 1970) and has been increasingly depicted with dynamic visualizations, such as animations (Fabrikant et al. 2008a). With static or small multiple visualizations, changes over time could be captured only implicitly by deriving changes of a sequence of static snapshots, whereas animations represent movement in a congruent and realistic way (Moellering 1976). The availability of a large amount of spatio-temporal data for a long period of time, e.g., movement data from GPS-enabled mobile devices, is supporting the development of novel visual analytics (VA) methods. In particularly, it promotes the increasing use of animations for representing and exploring movement data (Andrienko et al. 2010).

However, how viewers make decisions with and conceive of movement dynamics represented with animations is not adequately understood to date (Fabrikant & Lobben 2009; Shipley et al. 2013). Visuospatial decision-making processes with animations of movement data might be influenced by three main visual analytics (VA) dimensions: (1)

the use context (and the required decision-making task), (2) the animation display design, and (3) the user characteristics. User characteristics refer to viewer cognitive, perceptual, and affective processes involved in decision-making, e.g., spatial abilities, engagement, stress, motivation, expertise, age and gender. This thesis explores the interaction between the above-mentioned VA dimensions and their influence on specific decision-making tasks with animations.

1.1 Problem Statement and Motivation

This PhD thesis combines research in dynamic cartography, visual analytics, movement analysis, and cognitive science. It is part of a larger research project called "*Animated Visual Analytics of Movement*" (AniMOVE) that has been funded by the Swiss National Science Fund (Grant No. 200020-134646). This project extends the previous work done by Dr. Anna-Katharina Lautenschütz during her PhD project "*Visual Analytics of Spatio-Temporal Gaze Point Patterns in Eye Movements*" (Popeye). Her PhD thesis (2011) highlighted the need for future research to focus more on the cartographic aspects of visualization, e.g., how cartographers might make movement data displays visually more salient for users.

This work aims at empirically evaluating animated displays representing real-time movement data for effective and efficient decision-making. It is first intended to better understand how viewers perceive and conceive of movement patterns depicted with animations, as well as how viewers affectively react to specific animation designs. Second, it aims at constructing empirically validated design guidelines for perceptually salient, cognitively inspired and affectively engaging animations.

This research is motivated by the increasing attention in the GIScience community to represent movement data with appropriate dynamic VA displays, as a tool to facilitate the exploration and extraction of meaningful and relevant spatio-temporal information (Andrienko et al. 2010). Static maps are not completely suitable to depict spatio-temporal and dynamic phenomena, because of the difficulty of visualizing temporal aspects of a phenomenon. How to effectively include the time component in the visualization of spatio-temporal information has become a research priority in GIScience (Andrienko et al. 2007a). In this context, animations can be very useful in depicting the dynamics and behaviour of moving objects over time. Animations can support users in decision-making

and data exploration processes because of the congruent and coherent match with the user's mental representations (Tversky et al. 2002; Fabrikant et al. 2008a).

Moreover, in the last ten years, movement data at high spatio-temporal resolution has become increasingly available (Holyoak et al. 2008). This increasing movement data availability supports GIScience research and other research and application domains for the development of new VA methods for analysing and visualizing spatio-temporal phenomena (Andrienko et al. 2007a; Andrienko & Andrienko 2008). In air traffic control, examples of effective air traffic trajectory visualization systems aiming at facilitating the exploration of multiple aircraft trails are the FromDaDy system as well as interactive image-based information visualization systems (Hurter et al. 2009b; Hurter et al. 2014b).

Animation techniques are (or might potentially be) used to effectively analyse and visualize movement data, beyond GIScience applications, in such fields as air traffic control, crisis management, surveillance, sport analytics, movement ecology, transportation, and human health (Diamond et al. 2007; Hurter et al. 2009a; Klein et al. 2014; Dodge et al. 2016). Basically, in these research and application domains, animations might facilitate the achievement of two main tasks: adequate decision support for users working with real-time data and the effective exploration of spatio-temporal data. For example, users might need to promptly detect relevant movement patterns and movement changes of the dynamic phenomenon under study. They might also need to correctly recognise the behaviour of a specific moving object. Finally, they might be interested in better understanding the complex interactions between one moving entity and other entities, or between a moving object and its environment.

Despite this increasing interest in analysing and visualizing movement data with animated displays, many issues still remain under-researched. For example, only a few empirical studies on how users perceive and extract information from animated displays have been conducted in the last ten years (Fabrikant et al. 2008a). From these few empirical studies that have been published to date, it is difficult to derive any general conclusions that can help designers and cartographers to develop animations for efficient and effective user decision support (Kriglstein et al. 2012). Furthermore, a clear overview of how user perceptual, cognitive, and affective processes interrelate with each other, when they interact with a certain animation design, is still missing. Related to this subject, Kriglstein et al.

(2012) present a literature review to help researchers establish a more systematic overview of animation research.

Few empirical studies of animations especially address learning effects, presentations with interactive and non-interactive displays, and compare static and animated displays (Lowe & Boucheix 2010). Thus, they do not specifically focus on the effective visualization of spatio-temporal data for a specific use context, and on user characteristics influencing decision-making processes. For example, one finding is that users do not conceive of the movement dynamics of moving entities in isolation, but that user perception is influenced by the geographic context of the dynamic phenomena under study (Shipley et al. 2013). Further, few studies have focused on the depiction of real-time data for users' decision support with non-interactive animations, and on the empirical assessment of the interaction between animation design, context of use, and user characteristics. Few empirical studies have been executed with realistic decision-making tasks and in real-world use contexts (van Elzakker & Griffin 2013). In addition, only a few studies have investigated the interaction between user affective states and map design (Fabrikant et al. 2012). Affective states, such as motivation or stress, might also influence the efficiency and effectiveness of task performance and decision-making with animated displays. It is also unclear how display design choices (e.g., animation design types) might affect knowledge acquisition and inference making from animated displays (Fabrikant et al. 2008a), how this is influenced by the use context (e.g., type and complexity of the task for a specific use context), and how this is affected by user characteristics (e.g., affective state, spatial abilities, and expertise level).

For the above-mentioned reasons, this study aims to investigate more systematically the relevant factors affecting the effectiveness and efficiency of decision-making with non-interactive animated displays representing real-time movement data. We therefore propose a novel empirical research framework, based on cartographic design theories and supported by a series of empirical studies with users, to create perceptually salient, cognitively inspired, and affectively engaged animated displays of moving objects. This work also aims to derive empirically validated design recommendations to more generally support animation designers in answering the following question: How should animations of real-time movement data be designed considering the task and/or use contexts, and user characteristics? But this then leads to the question of, why is this question relevant in the context of Geography.

1.2 Relevance for Geography

Traditionally, movement data have been most often visualized by means of arrows or flow lines (Vasiliev 1997). In the past ten years, animations have been increasingly used in different cartographic application domains to represent spatio-temporal data as well (Andrienko et al. 2007b). Animations are widely used and are an important visualization means, e.g., to monitor dynamic processes and effectively detect spatiotemporal changes over time (Blok 2006). However, past and current GIScience research related to movement analysis has focused more on the development of computational methods for extracting relevant and meaningful information from these datasets (Klein et al. 2014). Limited empirical research has been conducted to date on the assessment of visualization methods and understanding user reasoning when decisions are supported by animations (Shipley et al. 2013).

Currently, there is also a lack of general cartographic design guidelines for animations (Fabrikant & Lobben 2009). Andrienko et al. (2010) point out that effective and efficient visual analytics tools can be generated only if computational and technical methods are appropriate to the perceptual and cognitive capabilities of the users. The need for GIScience and geovisualization to ameliorate and empirically assess this synergy between computational methods and the cognitive, perceptual, and affective capabilities of users has also been highlighted in previous international GIScience conferences (e.g., AGILE 2008, GIScience 2014, ICC 2015, etc.) and in the current research agenda of different cartography and VA associations is mentioned as a key research challenge (for example, the European coordination action VisMaster, http://www.vismaster.eu/, last access: 17.01.2017; and the ICA Commission on cognitive issues in geographic information visualization http://cogvis.icaci.org/, last access: 17.01.2017).

Previous empirical studies have often proposed a comparison between different map design types (mostly static versus dynamic map displays) and highlighted the supremacy of one design over another (Tversky et al. 2002). However, the challenge is not just to communicate how animated displays should be designed to effectively visualize dynamic information, but also to find an explanation of *why* a specific map design type works better than others from a cognitive, perceptual, and affective standpoint, and what factors might influence the effectiveness and efficiency of decision-making with animations (DiBiase et al. 1992; Koussoulakou & Kraak 1992; Harrower 2004, 2007; Slocum et al. 2004).

The purpose of this thesis is to fill the gap mentioned above with respect to the research framework and the empirical approach illustrated in the next section.

1.3 Research Workflow of this Thesis and Empirical Approach

This thesis has been developed according to an empirical research workflow consisting of four stages, as shown in **Errore. L'origine riferimento non è stata trovata.**. After a systematic investigation of relevant literature related to animation research, movement analysis, and cognitive psychology (i.e., Figure 2, box 1), an empirical research framework has been developed (i.e., box 2). This research framework focuses on an empirical assessment of animations by considering perceptual, cognitive, and affective mechanisms involved in user decision-making processes, and their use context. The empirical assessment consists of two user studies and is based on a methodological triangulation that couples eye movement data with electrodermal activity (EDA), electroencephalography (EEG), and questionnaires (i.e., box 3) (Maggi & Fabrikant 2014b).

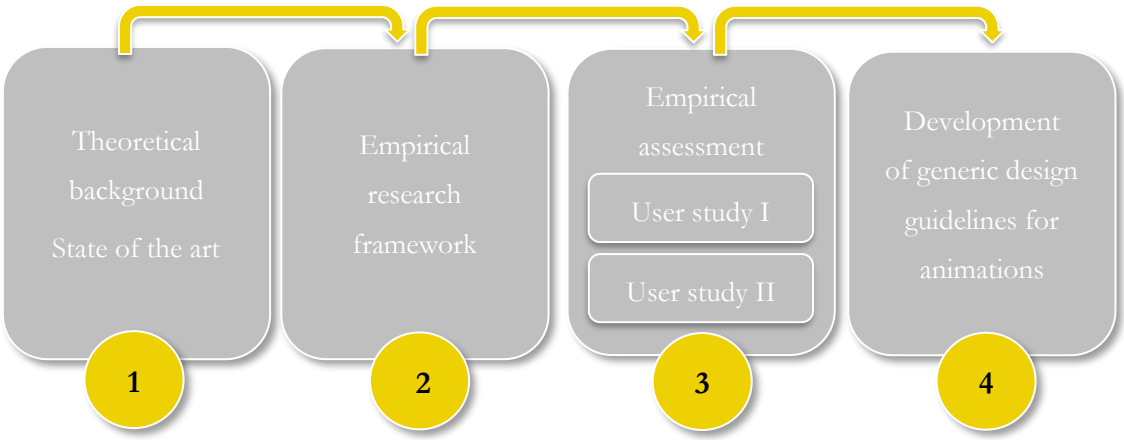


Figure 2: *Research workflow of the thesis including the user studies.*

In accord with current VA concepts (see http://www.vismaster.eu/faq/related-research-areas/, last access: 17.01.2017), we identified three main VA dimensions influencing decision-making with animations: the use context and task, the animation design, and user characteristics. As shown in Figure 3, these three VA dimensions interact with each other and influence user decision-making and task performance. The success of user decisions and task performance in a certain use context and for a specific task depends, in turn, on their cognitive, emotional, and perceptual processes.

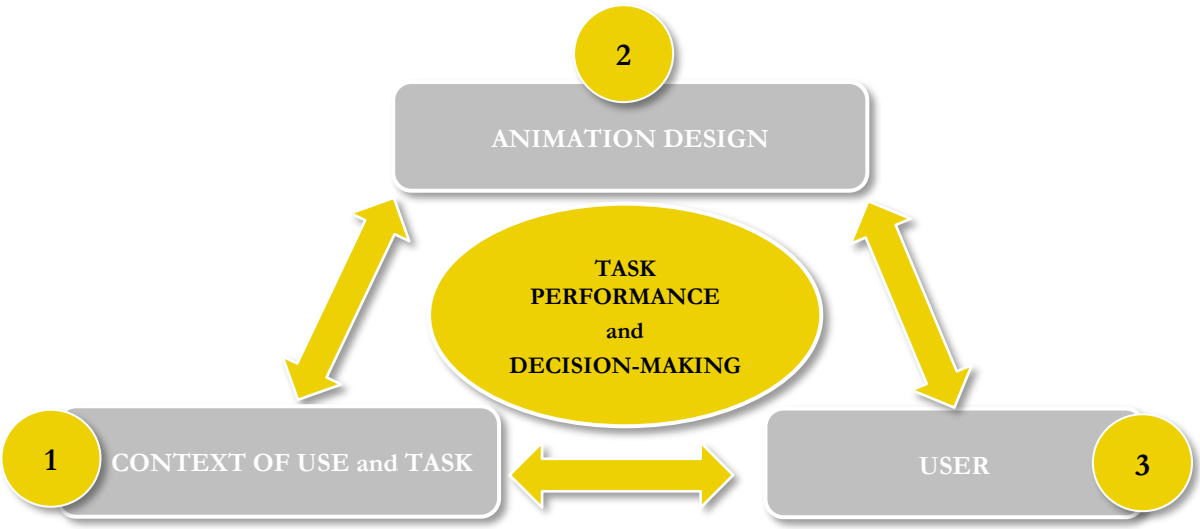


Figure 3: *Interaction between the three visual analytics (VA) dimensions (1) 'context of use and task', (2) 'animation design', and (3) 'user characteristics', influencing user decision-making with animations of movement data.*

In Figure 3, under the VA dimension *context of use and task* (i.e., box 1 in Figure 3), the following elements are relevant (especially in the domain of movement analysis and spatio-temporal data):

- **User domain**, such as surveillance, crisis management, and ecology research.
- **User task**, such as decision support or analysis support (e.g., movement change detection or prediction of future movement patterns of the moving objects under study).
- **Conceptual modelling of the movement and movement space**, such as the data or *moving object* (MO) type/number under study (e.g., two moving aircrafts), the relative *movement pattern type* (e.g., generic and behavioural movement patterns, such as two MOs that converge), the corresponding *movement parameters* (e.g., the speed and direction of the MOs under study), and the relative *movement space* in which the MOs under study is embedded (e.g., Lagrangian vs. Eulerian movement, constraints to movement, and continuous vs. discrete movement spaces) (Laube 2014).

Relevant factors of the VA dimension *animation design* (i.e., box 2) include:

- **Static and dynamic visual variables**, such as the rate of change of a display scene (or the smoothness of transitions between scenes), and the depiction of the

moving object trace (e.g., representation of past, current, and future object positions, or path history).

- **Interactivity** (not discussed in this thesis).

The VA dimension *user characteristics* (i.e., box 3) included the following factors:

- **Individual differences**, such as spatial ability, affective state (e.g., engagement, distress, and worry), and cognitive state (e.g., cognitive workload and motivation) of the user.
- **Group differences**, such as age, gender, and expertise (or also training and familiarity).

The two user studies were designed to empirically assess the interaction between the above-mentioned VA dimensions, i.e., animation design, context of use, and user characteristics (i.e., box 3 in **Errore. L'origine riferimento non è stata trovata.**). The statistical analyses and discussion of the obtained empirical results serve as the basis for the development of generic animation design guidelines (i.e., box 4 in **Errore. L'origine riferimento non è stata trovata.**).

This thesis is guided by the general research goals and research questions listed in the next two sections.

1.4 General Research Goals

This thesis pursues three main research goals. Following the three main factors influencing decision-making with animations mentioned in the previous section, they are as follows:

- **Focus on use/task context:** Identify *appropriate contexts of use and tasks to effectively evaluate animations of movement data.*
- **Focus on animation design:** Assess *how movement data might be effectively and efficiently represented* with animated displays (according to cartographic visual and dynamic variables) and develop *empirically validated design guidelines* for the construction of perceptually salient, affectively engaging, and cognitively inspired animations.
- **Focus on user:** Better understand *how movement dynamics is perceived, cognitively processed, and affectively sensed* by humans using animated displays for decision-making.

1.5 General Research Questions

To reach the above-mentioned general research goals, I proposed the following lead research question:

How should animations of real-time movement data be designed considering the task and/or use contexts, and user characteristics?

This research question has been divided into three research sub-questions (i.e., RQ 1, RQ 2 and RQ 3):

- **RQ 1 (USE/TASK CONTEXT):** How does the *use and task contexts* influence user visuospatial decision-making with animations depicting movement data?
- **RQ 2 (ANIMATION DESIGN):** Which *animation design characteristics* might be particularly useful to depict movement data for efficient and effective visuospatial decision-making with animations?
- **RQ 3 (USER):** How do viewer *perceptual, cognitive, and affective states* influence the effectiveness and efficiency of visuospatial decision-making with animations depicting movement data?

1.6 Structure of the Thesis

This thesis consists of eight chapters, including this introductory chapter. *Chapter 2* reviews the relevant literature and the state of the art concerning context of use, cartographic principles of the animation design, as well as user cognitive, perceptual, and affective mechanisms relevant for our empirical studies. *Chapter 3* presents the main motivation for the choice of the use case in the domain of air traffic control (ATC) and the results of ATC expert interviews. *Chapter 4* defines the methodology of the adopted empirical procedure for the two user studies. *Chapter 5* and *Chapter 6* describe the results and the corresponding statistical analysis of the two user studies. *Chapter 7* discusses critically the empirical results obtained from the two experiments and proposes general design guidelines for animations that are derived from the two user studies. *Chapters 8* summarizes the main contributions of this research work and presents a brief outlook on potential and future research areas.

CHAPTER TWO

RELATED WORK

This chapter presents prior work and the current state of the art and is organized according to the three main VA dimensions (i.e., use context/task, animation design, and user characteristics) influencing the effectiveness and efficiency of decision-making with animations (cf. Section 1.3 "*Research Workflow of this Thesis and Empirical Approach*"). In the first section of this chapter, some of the main research and application domains in which animations might be beneficial for decision support are described. In the second section, I present a literature review concerning animation design–related factors aimed at facilitating visuospatial information processing. In the third and last section, I review the most relevant user characteristics, including perceptual, cognitive, and affective mechanisms, influencing decision-making with animations.

2.1 Factors related to the Use, Task and Movement Data Context

According to Dodge et al. (2008), movement can be described by means of a set of movement parameters, such as speed (i.e., rate of position change), velocity (i.e., rate of position and direction change) and acceleration (i.e., rate of speed change) of a moving object. Further, movement can be categorized into two main movement pattern types: generic movement patterns, and behavioural movement patterns (Figure 4) (Dodge et al. 2008). Generic patterns are less complex than behavioural patterns and can be further distinguished into primitive and compound patterns. Primitive patterns are basic forms of movement patterns, such as 'full synchronization in time'. Full synchronization in time refers to movement changes (e.g., speed, direction and acceleration changes) occurring in space at the same time. Conversely, compound patterns are more complex than primitive patterns and involve a relation between several movement patterns, such as 'convergence'. Convergence is a specific kind of compound pattern that corresponds to several movement patterns, originated at different locations, but tending to a common location at the same time. Finally, behavioural patterns are patterns built by generic movement patterns, and are associated with a specific behaviour of the moving objects under study, such as a fighting behaviour among two or more animals.

However, in which context of use (i.e., the user research/application domain) and for what kind of task and type of movement pattern that animations are (potentially) advantageous for representing spatio-temporal information, compared to other visualization types, is still not clear (Kriglstein et al. 2012).

Animations are used and might be potentially helpful for decision support in different use contexts, e.g., transportation and human mobility, military and crisis management, surveillance, ecology, and animal behaviour research (Dodge 2015). In geographic research, for map-based decision-making Griffin et al. (2006) found that animations might be beneficial in helping users to effectively perceive and understand space-time patterns, and to track changes. Similarly, Slocum et al. (2004) found that animations are suitable for identifying general trends, whereas static small-multiples are more appropriate for comparing data from different times. Furthermore, animations are used as a standard display type in different application and research contexts. For example, animations are typically used in air traffic control (ATC), to visualize and monitor movement patterns in real time.

Regarding specific movement patterns where animations might be particularly suited, Gudmundsson et al. (2012) emphasised the appropriateness of animations to uncover speed patterns of single, moving objects, and to recognize converging movement patterns of a group of moving objects.

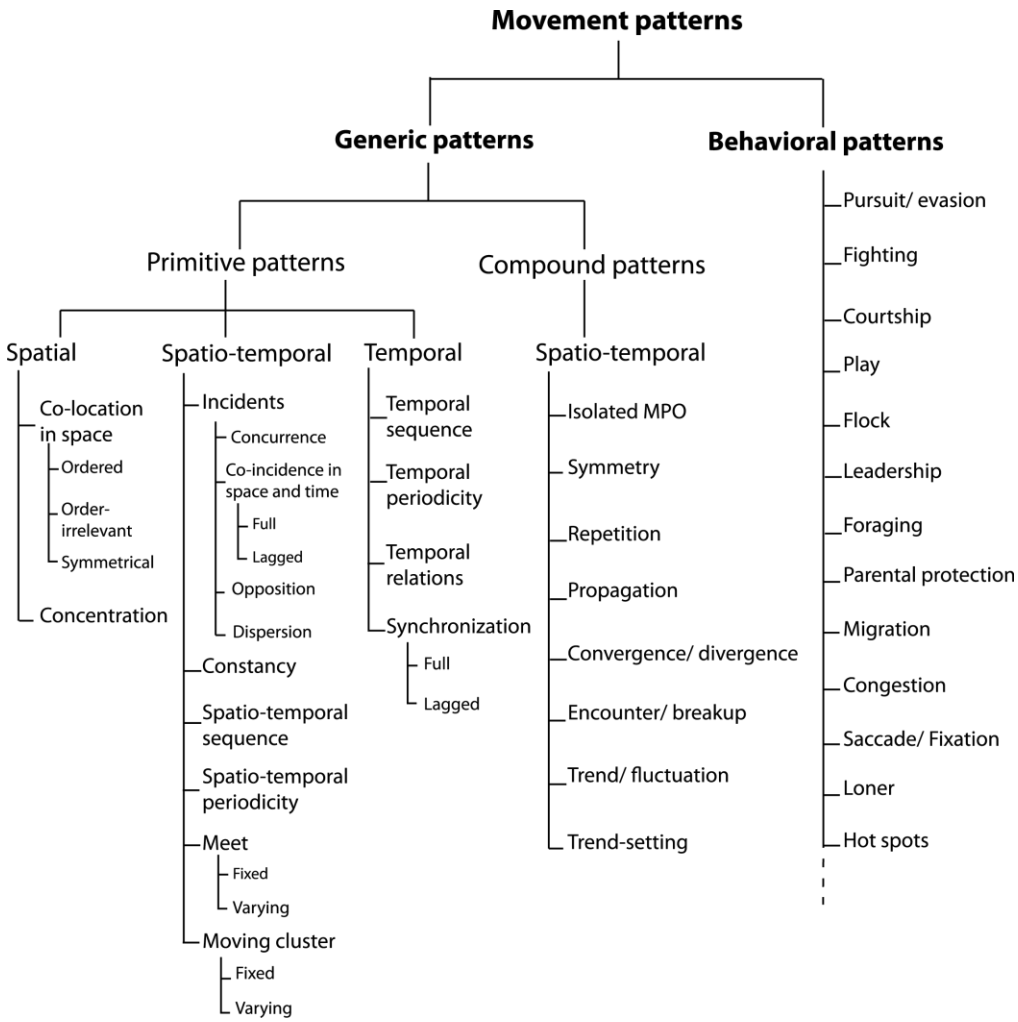


Figure 4: *Taxonomy of movement patterns (from Dodge et al. 2008).*

How animation design–related factors (e.g., rate of change of the display scenes) and user-related factors (e.g., expertise) might influence information extraction and decision-making processes involved with animated displays is explained in more detail in the next two sections.

2.2 Factors related to the Animation Design

2.2.1 Visual Analytics of Spatio-Temporal Data

According to Thomas and Cook (2005), visual analytics (VA) is defined as 'the science of analytical reasoning facilitated by interactive visual interfaces. People use visual analytics tools and techniques to synthesize information and derive insight from massive, dynamic, ambiguous, and often conflicting data; detect the expected and discover the unexpected; provide timely, defensible, and understandable assessments; and communicate assessment effectively for action'.

VA methods for spatio-temporal data modelling, representation and depiction have a long-standing research history. Traditionally, spatio-temporal phenomena have been represented statically by means of, e.g., maps organized in small multiples (Tufte 1990) and with the space-time cube (Hägerstrand 1970). To visualize spatio-temporal and movement data on static maps, cartographers often generalize or aggregate them, through appropriate methods across space or time, to avoid clutter (Andrienko & Andrienko 2011; Andrienko et al. 2011). For example, in the ATC domain, the edge-bundling technique has been developed for the analysis and visualization of small and large amounts of flight track data (Ersoy et al. 2011; Hurter 2016).

However, static visualizations prevent users from directly and intuitively perceiving the dynamics of moving objects, from conceiving their relationships between other objects, and from predicting future events, due to their inherent static nature (Shipley et al. 2013). For this reason, in the last ten years, the increasing availability of dynamic spatio-temporal data (e.g., movement tracking data from mobile devices) has induced an increasing interest in the GIScience community in depicting these data with novel visual analytics techniques. These methods aim to facilitate the identification and prediction of patterns and relationships between complex spatio-temporally coordinated events (Andrienko et al. 2010). In this context, animation seems to be well-suited to visualizing spatio-temporal data and the relationships among them effectively. However, how is the term 'animation' defined, and what kind of cartographic rules are especially relevant to effectively design animations for decision support? A possible answer to these questions will be found in the next section.

2.2.2 **Animation Design and Related Cartographic Rules**

The term animation comes from the Latin word animatio, which means 'to bring to life' or to 'give a soul' to something (Buziek et al. 2000). In data visualization and graphics, animations generate consecutive animated images to produce the perception of motion (Buziek et al. 2000). According to Bétrancourt and Tversky (2000), animations refers to 'any application which generates a series of frames, so that each frame appears as an alteration of the previous one, and where the sequence of frames is determined either by the designer or the user'. There are two different types of animations: recorded and real-time animations. Recorded animations refer to video or film, in which the single images are saved on a data medium after the generation of the animation and from there played and analysed. From the first cartographic animation created by Tobler (1970), technology for the purpose of creating cartographic animations of spatio-temporal data has changed rapidly, starting in the 1990s with Adobe Flash and progressing to the currently used JavaScript libraries for dynamic visualizations such as Torque by CartoDB, Processing.js, or D3 (Fish 2015).

Moreover, Tversky et al. (2002) highlights that effective and efficient graphics should conform to two principles: the Congruence Principle and the Apprehension Principle. The Congruence Principle states that 'the structure and content of the external representation should correspond to the desired structure and content of the internal representation'; graphics have thus to congruently match mental models of real-world phenomena. For example, animations can convey changes of spatio-temporal phenomena or events with frame changes over time in a consistent way. Consequently, animations should be an adequate graphic medium to represent spatio-temporal patterns and changes of moving objects (Fabrikant 2008). Tversky et al. (2002) contend that animations might be more suitable than static displays to depict 'real-time changes and reorientations in time and space', as well as 'qualitative aspects of motion or the microsteps, and the exact sequence and timing' of the spatio-temporal phenomena. However, very rarely has the empirical research in cartography addressed these aspects in their evaluations, and studies involving spatio-temporal data with potential users are rare (Kriglstein et al. 2012)

Previous empirical studies did not always find that animations had a clear superiority for visuospatial information extraction compared to other map display types, e.g., static displays. Often, past evaluations of animations produced controversial and contradictory

results. Comparing static displays with animations, Tversky et al. (2002) found that animations show information in a too fast and too complex manner to be effectively perceived and conceived. A possible explanation is that animations do not always conform to the Apprehension Principle, the second principle of good graphics. It specifies that 'external representations should be accurately perceived and appropriately conceived' (Tversky et al. 2002). Moreover, according to Shipley et al. (2013), relevant movement patterns in animations (Dodge et al. 2008) might be cluttered by other patterns that make it difficult for the users to identify them. Robertson et al. (2008) affirm that animations are useful to display small quantities of data. Previous studies demonstrated that users are able to track simultaneously a maximum of four objects with independent motions (Ware 2013). However, Cavanagh and Alvarez (2005) demonstrated that tracking more than four objects is possible by grouping similar behaviour of the moving objects. This is also in line with the Gestalt Principle of common fate (Koffka 1935), discussed later in Section 2.3.2, Cognitive Processes. In addition, past empirical studies comparing static with animated displays lack informational (e.g., information is equivalent in all the assessed displays) and computational (e.g., the ease of information extraction across the assessed displays is equivalent) equivalence (Fabrikant et al. 2008b). Often, animations convey more information than static displays, or allow interactivity (Schnotz et al. 1999). Finally, Harrower (2007) list different cognitive issues linked with animated displays, such as change blindness. Change blindness is a limitation of visual working memory capacity that prevents users from perceiving relevant changes within dynamic displays (Goldsberry & Battersby 2009).

Cartographers have developed appropriate design and cartographic rules to effectively and efficiently perceive and conceive animations, and thus to overcome the cognitive limitations of users as described above. Bertin (1967) proposed design guidelines for effectively communicating visual information on graphics. He described seven basic visual variables, i.e., position, size, shape, colour value, colour hue, orientation, and texture, to support cartographers in designing adequate information visualizations. Later, Morrison (1974) proposed additional variables, such as colour saturation, crispness, and transparency. Even if Bertin's visual variables were defined for static maps and not for animated displays, DiBiase et al. (1992) and MacEachren (1995) demonstrate that they might be applicable for animations, as well. To adequately communicate spatio-temporal information with dynamic visualizations, a set of dynamic visual variables have been defined, i.e., display

moment, scene duration, scene frequency, frame order, rate of change between scenes, and synchronization of spatio-temporal phenomena (DiBiase et al. 1992; MacEachren 1995). However, a clear definition of the concept of dynamic visual variables is currently still missing. Ben Rebah and Zanin (2011) propose additional dynamic visual variables, i.e., the change rhythm, the proportionality variable, and the trajectory variable. According to DiBiase et al. (1992) and Griffin et al. (2006), an important aspect for effectively designing animations and perceiving apparent motion is the interaction of three factors: the distance an object moves, the stimulus duration, and the frame rate.

A third design element that cartographers might manipulate is animated transitions (e.g., Chevalier et al. 2010; Shanmugasundaram et al. 2007; Robertson et al. 2008). Animated transitions, or tweening, between scenes are useful for detecting small changes in data, because they allow the anticipation of spatio-temporal changes and can prevent attentional blindness effects (Fabrikant et al. 2008a). The question is to what extent tweening is useful; if changes are too smooth, it might be difficult for users to effectively perceive them. Animated transitions are related with the dynamic variable 'rate of change' and they are discussed in more detail in the next section.

Even if some solutions exist to design animations effectively, there is still a lack of conceptual framework for the creation of effective animations (Lautenschütz 2011). In addition, for real-time data, when spatio-temporal information has to be visualized according to the real movement dynamics of the depicted entities, the solutions proposed above from past research might be able to be used only to a limited extent. As Fish (2015) argues, cartographers are constrained by the spatio-temporal dimension of the data. She further claims that the representation of spatio-temporal data in animations does not occur ideally according to the traditional graphic rules of animators, such as the 12 rules of animation of Walt Disney (Lasseter 1987), because spatio-temporal information has to be depicted inherently to the reality. The question that arises is how cartographers might depict spatio-temporal data inherently to the reality without compromising the perceptual, affective, and cognitive capabilities of the users. Furthermore, Kriglstein et al. (2012) affirmed that cartographers need evidence-based design guidelines for efficient and effective animations that clearly describe how animations should be designed, delimit the context in which animations should be used, and expose which kind of features in animations are particularly helpful for users. The next section presents in more detail the difference between semi-static and continuous animations with respect to the rate of

change (or smoothness of transitions) of the display scenes and the depiction of path history.

2.2.3 Rate of Change and Path History in Animations: Semi-static versus Continuous Animations

As highlighted in the previous section, in past research in cartography static and dynamic representations have often been sharply distinguished and evaluated by comparing one against the other. To overcome this sharp distinction, Nossum (2013) proposed a new kind of dynamic visualization to depict spatio-temporal information, i.e., the *semi-static animation*. Semi-static animations are dynamic representations that combine the qualities of both static and animated displays. Semi-static animations do not depict spatio-temporal data continuously, but rather like a sequence of static displays. In addition, spatio-temporal information is visually available at any time, i.e., in which past, present, and future information (i.e., *path history*) is displayed when running the animation. Figure 5 shows an example of a semi-static weather forecast animation map, in which the corresponding past, present and future weather forecast symbols are visualized by means of a horizontal moving bar (i.e., path history) for the whole duration of the animation.

In cartography, semi-static and continuous animations can be distinguished by means of the dynamic variable *rate of change*, or the smoothness of the transitions between display scenes. Transitions between display scenes in semi-static animations occur abruptly (i.e., non-tweened, or non-gradual and non-continuous, changes of the depicted phenomenon), while transitions between display scenes in continuous animations occur smoothly (i.e., tweened, or gradual and continuous, changes of the depicted phenomenon) (Battersby & Goldsberry 2010).

Conversely to semi-static animations, continuous animations convey spatio-temporal phenomena in a direct and realistic manner, due to their smooth and uninterrupted visuospatial changes over time. Real-time and continuous animations can be produced and analysed from a sequence of static images with a frame rate of 24 Hz (i.e., 24 images per second) (Harrower & Fabrikant 2008). This frame rate corresponds to the industry standard speed of TV and cinema movies (even if recently this value has reached 48 Hz in some cases) and is the minimum value to effectively perceive motion (Buziek et al. 2000).

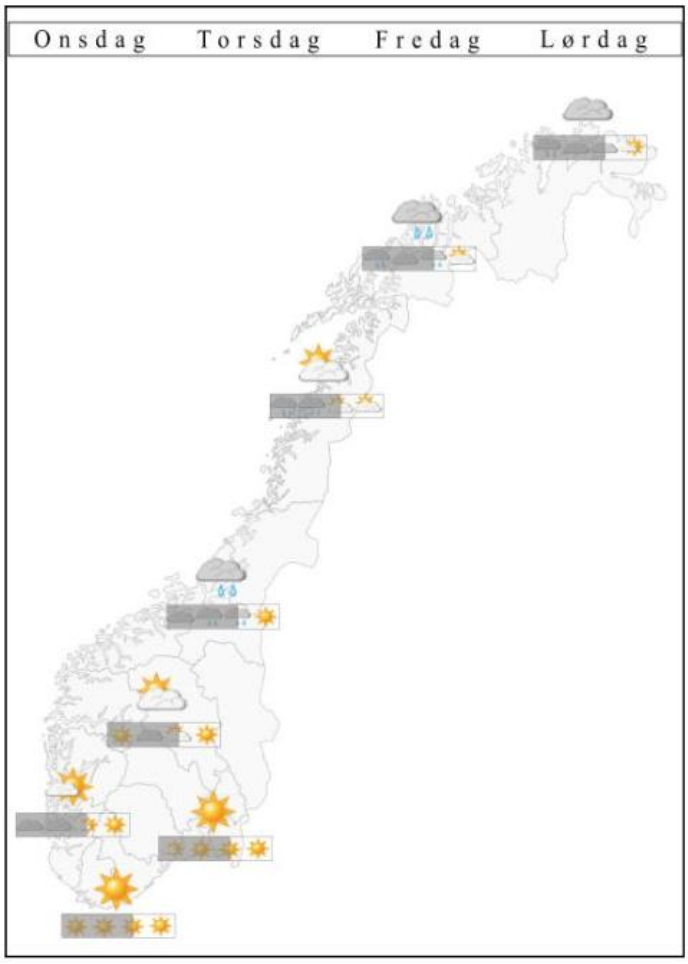


Figure 5: *Example of a semi-static animation depicting past, present and future weather conditions over four days in Norway (from Nossum 2013, http://geomatikk.ntnu.no/projects/semistatic, last access: 17.01.2017).*

However, research on dynamic visualization of movement data has mainly focused on the computational and technical aspects of extracting relevant information from very large data sets, and less on the implications of map designs on user decision-making processes. Only very few efforts have been made to conduct empirical research of real-time data animations for decision support. Andrienko et al. (2010) claim that cognitive and usability issues have to be discussed more to improve animated VA tools. Moreover, Fabrikant and Lobben (2009) point out that further empirical studies are needed in order to assess the effectiveness and efficiency of dynamic VA visualizations by considering the perceptual, affective and cognitive issues of users.

Map animations have profoundly changed the way cartographers depict, and users process, visuospatial information (Buziek et al. 2000). The way in which also user-related factors (including perceptual, affective and cognitive mechanisms) influence information processing with animations is the topic of the next sections.

2.3 Factors related to the User

van Elzakker and Griffin (2013) emphasize that the current research in geovisualization and cartography needs to focus more on users and on the development and implementation of user-centred design methods. Not only do map design, map purpose, and map context play an important role in user information inferences and decision-making processes with dynamic displays, but user backgrounds and training are also relevant in this context (Lloyd & Bunch 2010). User-related factors refer to both individual and group differences across users that might influence the way they understand the graphic interface (Slocum et al. 2001; Roth 2013). However, the implications of individual and group differences in visuospatial decision-making with animated displays and more generally with cartographic visualizations have only been tested with users sporadically to date (Hegarty & Waller 2005; Montello et al. 1999; Slocum et al. 2001; Wilkening & Fabrikant 2011).

Human visual attention and visuospatial inference with graphic displays may be modulated by three main mental processes: perceptual, cognitive, and affective processes. Previous empirical work in cartography focused their investigations mainly on the first two, i.e., perceptual and cognitive processes. These two mental mechanisms are typically described by psychologists as bottom-up and top-down processes. They are explained in more detail in the next two sections. The characterization and analysis of affective and emotional mechanisms influencing map-related decision-making has been typically neglected in the past as well as in the current cartographic and GIScience literature. However, according to the Yerkes and Dodson law (1908) (also known as the Inverted-U Model), performance related to specific map-related tasks and use contexts might be influenced by the affective state of the users, inclusive positive states (e.g., motivation) or negative states (e.g., stress).

Moreover, there is a wide range of potential differences between individuals that might influence human interaction with animations and graphic displays. Users with dissimilar abilities might interact with a graphical display differently when they are performing the same task. User factors influencing visuospatial inference with animations, and more generally with map displays, can be divided into two main groups: user individual characteristics and group differences. Individual characteristics of particular relevance in user-centred design research are, for instance, perceptual and cognitive skills, spatial

abilities, and affective states (including the attitudes and preferences of users). Group differences concern differences across groups of individuals. For example, in cartographic research, user expertise or training (novice vs. expert), gender (female vs. male), and age (young vs. old) are often mentioned as relevant group differences. In cartographic research, aside from user spatial abilities, group differences across users interacting with graphical displays have been analysed more than individual differences. There are also other factors, such as the cultural background and education of the users, that might be relevant in decision-making with animations, and more generally for cartographic interfaces (Slocum et al. 2001).

In the next sections, I review and discuss the most relevant perceptual, cognitive, and affective processes, as well as individual and group differences, influencing visuospatial interaction with animations by users.

2.3.1 **Perceptual Processes**

Bottom-up processes modulate mental mechanisms that are carried out directly by external visuospatial stimuli. The visual information is perceived through the human eyes and successively elaborated by the brain as an image, where it is identified and recognized (bottom-up, or from perception to cognition).

Motion is a very powerful salient visual cue, because it can attract user attention very easily and quickly compared to static displays (Petersen & Dugas 1972). Motion is one of the strongest pre-attentive features (inclusive colour, orientation, size, blinking and changes in animated scenes) that allows users to easily identify a target object among a set of other visual distractors (Ware 2013). In animations, the saliency of a certain moving object, or a group of moving objects with similar behaviour, is determined by the dynamic contrast strength with its visual background or context (Boucheix & Lowe 2010). Similarly, anomalous or unexpected movement patterns occurring on a homogeneous display background (e.g., an object moving faster or in a different direction compared to other objects) become perceptually salient, and thus easily discernible, compared to their context. The distinctiveness of a visual cue depends on the number of target objects among other graphic symbols, as well. If the target object is the only one on the screen that is distinctive compared with the context (e.g., an aircraft moving with a different orientation compared

to other aircraft), it pops up very quickly in a display and it is thus easily identified by the user.

Itti and Koch (2001) developed a computational model of human visual attention in which bottom-up attentional processes are particularity emphasized. This pre-attentive vision model predicts the focus of the visual attention based on perceptual salient information. Similarly, in air traffic control (ATC) research, the Noticing-Salience, Expectancy, Effort, and Value (NSEEV) model has been developed to systematically assess bottom-up (i.e., visual salience of events) as well as top-down processes (i.e., user cognitive effort, expectancy of an event occurrence and value/importance of task/event) in ATC tasks and dynamic visualizations (Steelman et al. 2011). This model emphasizes the importance of motion cues in modulating visual search and visual attention in ATC tasks. For example, by animating a visual scene, air traffic operators are effectively attracted by dynamic salient information, such as animated notification design types (Imbert et al. 2014).

However, motion is not only a powerful visual cue that effectively captures human visual attention, it also allows users to perceive the animacy and causality of the depicted motion features. Michotte (1963) claimed that humans perceive immediately and directly (the illusion of) causal relationships between moving objects in a simple animation and under certain spatiotemporal conditions, i.e., the relative velocities and proximity of the moving objects, and the timing of the motion change. Previous studies in the context of dynamic data visualization emphasized the value of causal animation for superior performances during learning and memory tasks compared with static displays and non-causal animations (Irani & Ware 2003; Kadaba et al. 2007). Further, Michotte (1963) also suggested that moving objects in a dynamic scene might be perceived as being alive. As a consequence, consistent with Gaur and Scassellati (2006), this allows users not only to identify movement patterns, but also to detect intentionality and specific behaviours (e.g., such as fighting), goals (e.g., 'trying to get over here'), mental states (e.g., 'wanting to get over here'), personality traits (e.g., shyness), and emotions (e.g., anger) of the depicted moving objects, as shown in Figure 6. For example, Stewart (1982) demonstrated that the perception of animacy might be modulated by three types of motion: start from rest, changes in direction to avoid collision, and direct movement towards a goal.

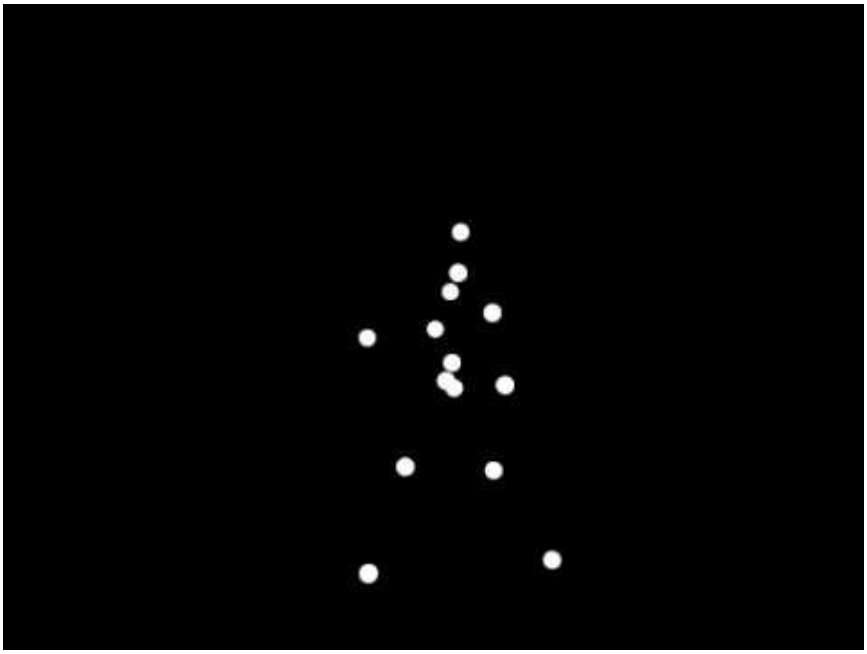


Figure 6: *Male-point light walker: An example of how animation might elicit in the users the illusion of animacy of the depicted object (Animation can be seen at: https://www.youtube.com/watch?v=r0kLC-pridI, last access: 18.01.2017).*

Further, previous empirical studies with graphic displays and dynamic visualizations highlight the importance of the user's prior knowledge, training, or expertise related to information processing (Gonzalez 1996; Kriz & Hegarty 2007; Wai et al. 2009). For example, different studies demonstrated that expert users performed visuospatial tasks more accurately than novices, due to their acquired prior knowledge and training (Fabrikant et al. 2010; Kalyuga et al. 2003; Lloyd & Bunch 2005; Wright et al. 2008). According to prior studies on animations (Boucheix & Lowe 2010; Fabrikant & Goldsberry 2005; Kriz & Hegarty 2007; Lowe 1999) this might be due to the influence of users' prior knowledge on recognizing thematically relevant information compared to perceptual salient information. The findings of these studies underlined that novices are often visually attracted by perceptually salient information, because they attend to the visual information mainly bottom-up. In contrast, information processing by experts is predominantly guided by top-down mechanisms.

Lack of training and familiarity with the tested animated displays and with the specific task might have a negative effect on performance, because novices do not possess the necessary mental models to effectively perceive the visuospatial information depicted in animations. In fact, participants without prior knowledge relative to a specific task and to the subject under study, have difficulties in mentally conceptualizing the presented

information and so appropriately solving the task (Lowe & Schnotz 2008). This has also been demonstrated in studies in the context of psychology and sport disciplines with expert trainers and with non-experts (Khacharem et al. 2013). As introduced before, user visual attention might be modulated by changes at event boundaries. However, the segmentation of a specific action might occur according to bottom-up or top-down principles, depending on the user's prior knowledge and familiarity with the task and context (Zacks et al. 2007). Novices segment events mainly according to fine-grained, low-level movement pattern changes (i.e., according to bottom-up mechanisms), whereas experts segment events more coarsely by considering the intentions and goals of the moving objects (i.e., according to top-down mechanisms) (Zacks & Tversky 2001).

2.3.2 **Cognitive Processes**

Human attention can be directed by top-down processes as well. In opposition to bottom-up processes discussed in the previous section, top-down processes are typically induced by higher cognitive mechanisms. In this case, human visual attention is guided by factors such as specific goals, targets, expectance of an event, importance/priority of a task, cognitive workload capacity, and expertise (also referred to as a user's prior knowledge, training, or familiarity).

The *Gestalt Theory* highlights the importance of top-down mechanisms and holistic learning in visual information processing (Koffka 1935). During a lecture at the Kant Society in Berlin in 1924, Wertheimer declared: 'There are contexts, in which what is happening in the whole cannot be deduced from the characteristics of the separate pieces, but conversely; what happens to a part of the whole is, in clear-cut cases, determined by the laws of the inner structure of its whole' (Krapp 2005). Therefore, visual cues are processed by viewers not as single units in isolation, but globally, as a whole entity, embedded in their specific visual context. In the context of animated displays, the *Gestalt principle of common fate* (Koffka 1935) is of particular relevance, because it demonstrates that users process dynamical visual information by grouping similar movement patterns (e.g., objects moving with similar speed or in the same direction).

Furthermore, the interaction between information processing with spatio-temporal objects and top-down mechanisms introduces another theory related to events and event perception, the *Event Perception Theory* (Shipley & Zacks 2008). This theory argues that

humans perceive continuous and dynamic external information as discrete entities and well-defined spatio-temporal units. According to Zacks (2008), event perception is the set of cognitive mechanisms by which observers pick out meaningful spatio-temporal wholes (see mechanisms of Gestalt grouping discussed earlier) from the stream of experience, recognize them, and identify their characteristics. Interaction between moving objects and human actions are examples of perceptual events. Human actions implicate different typologies of interactions, e.g., the interactions among people, or interactions between humans and the environment, objects, or animals, etc. Usually, human actions are followed or caused by previous events, are directed to specific goals, and precede future actions. Humans typically segment actions at their corresponding event boundaries, i.e., the locations, or time intervals, where/when the movement of the considered object changes in the space and time dimension, or where/when the object changes its characteristics or attributes. For example, in an experiment with animations showing a soccer player's actions, participants recognized and segmented the player's actions by ball change possession (Huff et al. 2012). Physical or conceptual changes determine the start and the end of an event. These changes at event boundaries illustrated with animations are thus very important in guiding and attracting user visual attention over time (Huff et al. 2012).

However, how users segment events and, more in general how users process visual information on and make decisions with dynamic displays, depends on their individual characteristics and group differences. The most relevant individual characteristics for this thesis are the following: self-confidence, satisfaction and preference of users for using a certain map type, and spatial skills.

Concerning *user preferences* for different display design types, previous empirical work with animated displays suggests that users often judge dynamic representations as more enjoyable, helpful, and exciting for visuospatial information processing in comparison with static displays (Kriglstein et al. 2012). According to the participant responses to the post-test questionnaires of our experiments, participants reported that they prefer the display design type with which they feel themselves more confident, and that they thus believe they have solved the required task more easily. Preference ratings are thus related to self-confidence and self-perceived ease of use. Furthermore, self-confidence in visuospatial decision-making and map-reading tasks is perceived differently between females and males. Males often tend to overestimate their own abilities and task performance. Conversely,

females frequently express lower self-confidence in various tasks including visual ones (Lenney 1977; Wilkening & Fabrikant 2011).

Studying human *spatial abilities* in processing visuospatial information is important in many fields and activities, such as navigating in a city or a ship in the ocean, finding efficient routes to and from places, running a triathlon competition, etc. Further, Wai et al. (2009) advocate that human spatial skills might play a relevant role in predicting future achievements in science, technology, engineering, and mathematics (STEM) domains. To measure human spatial abilities, researchers have created various psychometric tests with the objective of solving problems with visuospatial figures that involve visual perception (Ekstrom et al. 1976; Carroll 1993). There are five visuospatial factors affecting human skills related to visual perception: visualization, spatial relations, closure speed, flexibility of closure, and perceptual speed (Carroll 1993). Some standard psychometric tests, such as Hidden Patterns (French et al. 1963), Card Rotations (French et al. 1963), and Vandenberg Mental Rotations (Vandenberg & Kuse 1978) tests, evaluate both static and dynamic spatial abilities of people, as well as 2D and 3D visualization parameters (Montello et al. 1999). Consequently, some of these tests are particularity suited for measuring user spatial abilities with dynamic visualization.

Group differences that are relevant for this thesis are the following: gender, age and expertise. Various empirical studies and theories demonstrate that *gender* is a relevant user-related factor influencing map-reading and decision-making with graphical displays. The *Evolutionary Theory*, as well as the *Hunter Hypothesis*, state that human spatial and verbal abilities are innately different between females and males, because of their different evolutionary adaptations (i.e., due to different environmental and social pressures between females and males) (Choi & Silverman 2003; Lloyd & Bunch 2008). This difference in gender might also have a consequence on female and male cognitive and spatial skills (Weiss et al. 2003). This relationship might have a significant influence in traditional psychometric paper-and-pencil tests, especially in the mental rotation test, but it is still not clear how gender differences influence more complex and larger-scale visuospatial tasks, such as map reading or way-finding (Montello et al. 1999). Information processing with maps is difficult to generalize and often depends on the map type and map task (Montello et al. 1999). In any case, the extensive literature concerning gender-related cognitive and spatial abilities shows that males perform better in spatial tasks, e.g., in mental rotation and map reading (Zinser et al. 2004), whereas females are mostly better in tasks involving verbal

abilities and in object location (Neave et al. 2005). Furthermore, previous studies on dynamic spatial reasoning tasks, e.g., in estimating the relative speed of moving objects depicted on animated displays, demonstrate that males perform better than females (Law et al. 1993; Montello et al. 1999).

Gender differences also influence *human brain activity*. Previous studies on brain lateralization (Sun & Walsh 2006), i.e., brain activity predominance in the left or in the right hemisphere, indicate that females and males activate different brain regions when they are processing information on graphic displays. Annett's *Right-Shift Theory* (Annett 2002) points out that people with a higher right hemisphere activity prefer to process verbal information to solve problems and complete visuospatial tasks, whereas people with a higher left hemisphere activity prefer to use spatial information (Lloyd & Bunch 2008). This, in turn, is again correlated with differences in gender and in the spatial abilities of females and males (Clements-Stephens et al. 2009; Weiss et al. 2003).

Age is also a factor influencing human cognitive and spatial abilities. Driscoll et al. (2005) argue that aging negatively influences human cognitive and spatial abilities. Aging is critical in affecting working memory, as emphasized in different studies, e.g., Salthouse (1990). Further, Salthouse (2009) affirms that cognitive abilities already begin to decline between an age of 20 and 30 years.

However, despite this multitude of empirical studies on the effect that individual and group differences might have on human visuospatial processes with graphic displays, it is still not clear to which extent these differences influence understanding and decision-making with animated displays, and how these user-related factors are interconnected with the animation design type employed and the complexity of the depicted visuospatial information (Fabrikant 2005; Maggi et al. 2016). Bunch and Lloyd (2006) discussed how the *Cognitive Load Theory* (CLT) might be applied in the context of geovisualization to effectively assess human information processing with maps. According to this theory, the success of user decisions depends on the animation design type, which, in turn, depends on user cognitive processes and cognitive load. User *cognitive load* is defined as 'the amount of work needed to acquire and use information' (Bunch & Lloyd 2006). So, user cognitive load might be high or low with respect to their capacity to process information more or less efficiently and effectively.

According to Bunch and Lloyd (2006), user cognitive load may be influenced by three main factors: the animation design, the task complexity, and user characteristics (i.e., individual and group differences), as shown in Figure 7. Mental load (i.e., demand of the map-related task and design), mental effort (i.e., the actual cognitive capacity required from a specific task), and task performance (i.e., the combination of mental load and mental effort with respect to the specific task and map display) are three assessment parameters related to cognitive load that can be measured during experiments.

Cognitive load is associated with (limited) working memory and long-term memory capacity (Bunch & Lloyd 2006). Two user characteristics might affect these two types of memory. Spatial abilities influence working memory, expertise affects long-term memory. According to Hegarty and Waller (2005), low- and high-spatial individuals create a mental representation of a graphic display qualitatively differently, where low-spatial users mentally retain visuospatial information with more difficulty than high-spatial users. Further, the same spatio-temporal information might be depicted with a simple or a more complex display design. Users with different expertise levels process this information differently with the two design types (Figure 8). Novices deal with complex maps, with a higher cognitive load than experts. Conversely, experts might need to solve difficult tasks that require complex maps in order to deeply understand specific spatio-temporal events or relationships among data (the *expertise reversal effect* of Kalyuga et al. 2003). As Kriz and Hegarty (2007) point out, processing continuous spatio-temporal information with animations involves complex interrelationships among bottom-up and top-down mental mechanisms. Temporal constraints intrinsic to animations can cause an increase of working memory and thus of cognitive load, compared to static displays (Mayer et al. 2005). In contrast, top-down processes, e.g., expertise and expectations of the user may positively influence task performance in map-related tasks by allowing chunking of information (from long-term memory), which in turn reduces cognitive load (Gobet 2005).

The interaction of users with a particular graphic interface, in particular with animations, not only influences their task performance, decision-making and preferences, but it also impacts and elicits different *users' emotions and their affective states*. Emotional and cognitive states influence each other reciprocally. This interaction between cognitive, emotional and affective states is explained in more detail in the next section.

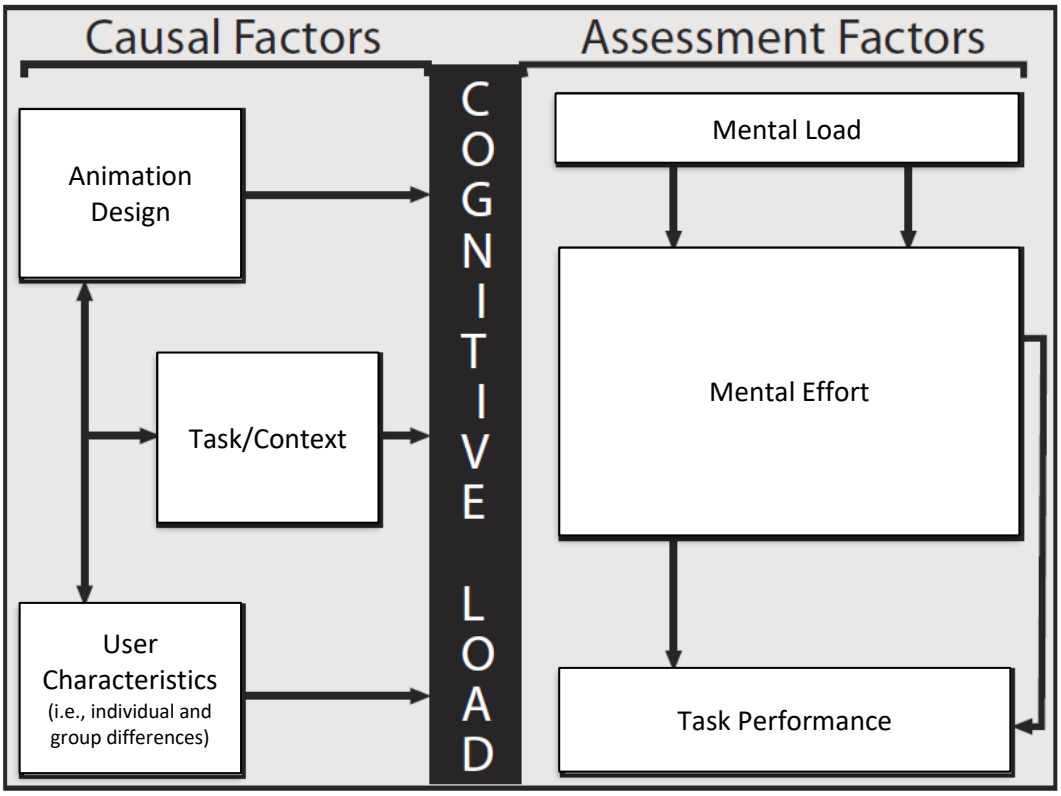


Figure 7: *Factors influencing the cognitive load of map users (modified from Bunch & Lloyd 2006).*

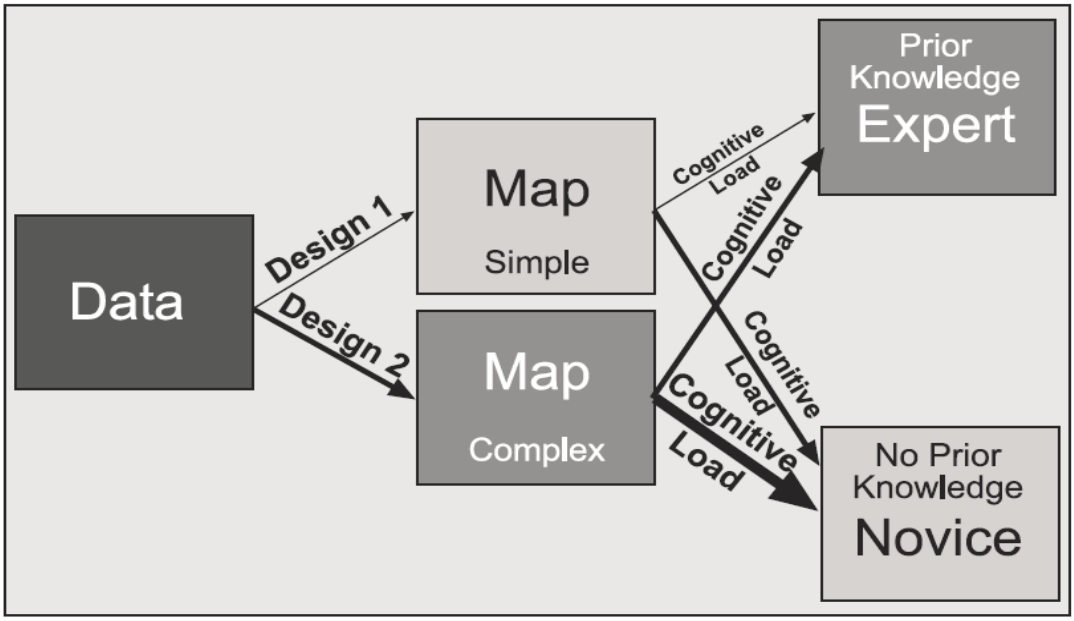


Figure 8: *The effect of map design, information complexity, and expertise in map-reading tasks (from Bunch & Lloyd 2006).*

2.3.3 Emotional and Affective Processes

When users make decisions by using a certain graphical interface, they might be happy, engaged, and calm, or, by contrast, frustrated, distressed, or worried. Further, the use of a particular graphic interface might activate some pleasant or familiar memories that induce positive emotions and enhance motivation. Conversely, the representation of visuospatial information with a certain design type might trigger negative psycho-physiological states due to internal (e.g., negative memories or stress) or external (i.e., task or data complexity) factors. Furthermore, a user's negative affective and emotional state during a map-based task can also be interpreted as an indicator of cognitive overload.

Theories of *Embodied Cognition* claim that cognitive processes are strongly influenced by embodied interactions with the environment, and vice versa (Wilson & Golonka 2013). Human cognition thus interacts continuously and directly with human emotions and human affective states. For example, user emotions and their affective states strongly influence goal-oriented cognition and information searching.

However, there is no exact definition and categorization of human emotions, to date. For example, Hockenbury & Hockenbury (2011) define an *emotion* as 'a complex psychological state that involves three distinct components: a subjective experience, a physiological response, and a behavioural or expressive response'. Ekman et al. (1972) identify six basic and universal emotions: fear, disgust, anger, surprise, happiness, and sadness. Plutchik (2001) distinguishes between eight kinds of emotions in his *Wheel of Emotions Theory* related to colours: joy, sadness, trust, disgust, fear, anger, surprise, and anticipation.

Positive and negative emotions influence task performance. According to the *Approach–Withdrawal Hypothesis* (Davidson et al. 1990), task performance can be measured and analysed in brain activity. In previous empirical studies (van Dantzig et al. 2008), in approach conditions (i.e., implicating positive emotions) users are faster and respond correctly compared with a withdrawal situation (i.e., implicating negative emotions).

The valence of emotions (i.e., if positive or negative) thus influences cognitive processes. However, more recent studies in neuroscience suggest that also *motivation* (or motivational intensity) might influence cognitive processes, such as task performance and decision-making (Harmon-Jones et al. 2012). Motivational intensity is a type of affective state and is defined as the impulse to move toward or away from a stimulus (Harmon-

Jones et al. 2012). Motivation is thus similar to emotion, but it is more driven to goal fulfilments and actions to obtain 'reward', or desirable reference values, and to avoid 'punishment', or undesirable reference values (Pessoa 2009; Roseman 2014). Harmon-Jones et al. (2012) point out that affective states with high motivation intensities narrow cognitive processes, whereas affective states with low motivation intensities broaden cognitive processes.

The usability degree and design type of graphical interfaces might influence user performances as well. According to the *Cognitive Fit Theory* (Vessey & Galletta 1991), the design of graphical interfaces has to fit the cognitive challenges of the users, e.g., goals, cognitive workload and emotions (e.g., anxiety or motivation) (Fraser et al. 2015). For example, if the complexity of the depicted information, or task, increases, the user's cognitive load increases. At the same time, the user might become more stressed, less motivated, experiencing increased heart rate and decreased pupil size (Brunken et al. 2003). For this reason, measuring the affective states of users, as explained in Chapter 4, *Methodology*, might be used as an additional measure, beyond standard usability metrics, to compare the effectiveness and efficiency of certain display design types. More specifically, how user *stress levels and motivation* changes by solving a certain map-related task might affect their task performance.

However, despite the interesting and useful aspects that the analysis of emotions and affective states might have in the context of geovisualization and cartography (i.e., in studying the usability of specific display design types), very little empirical research has been done to date (Griffin & McQuoid 2012; Klettner & Gartner 2012). Further, only a few research efforts have been made to investigate the implications of emotions in map-related decisions and user performance in solving visuospatial tasks (Fabrikant et al. 2012; Maggi & Fabrikant 2014a). Some of the existing examples are described below.

Emotional Cartography visualizes human emotions and studies the implications that these visualizations might have from political, social, and cultural points of view (Nold 2009). In the 17th Century, *The Carte du Tendre* (Madeleine de Scudéry 1654) was created with the goal to depict the geography of love. On this map, different aspects of the love emotions have been traced into an allegorical space of roads, rivers, mountains, and villages. In recent years, the Greenwich Emotion Map was created by local residents between 2005 and 2006 to visualize the relationship between space and emotions at a specific moment

(Nold 2009). Further, some researchers (Zeile et al. 2009; Gartner 2012; Huang et al. 2014) tracked human emotions with a sensory device and GPS to locate and categorize emotions (e.g., anger, calm, etc.) by walking through a city. This might contribute to optimizing urban planning and to the development of 'smart cities' (Zeile et al. 2009). Furthermore, regarding map designs and emotions, Fabrikant et al. (2012) presented a novel approach to assess the aesthetics of different map designs and the effect of certain colour combinations within maps on users' emotions.

In previous studies, where static displays were compared with animations, participants often reported that they enjoyed and preferred solving a task with animations compared with static displays, even if they did not perform better than with static displays (Rieber 1991). Reiber found that animations not only effectively influenced the learning process compared to static displays, but that animations had a positive effect on users' *continuing or intrinsic motivation* as well. Participants may be more motivated to solve a task with animations than with static displays. User motivation may considerably impact their interaction with cartographic interfaces, and consequently their task performance (Roth 2013).

The *Yerkes-Dodson law* (or also called *Inverted-U Model*) (Yerkes & Dodson 1908) shows the relationship between a user's task performance and their physiological arousal intensity. At the same time, their arousal intensity is linked with their stress level and motivational intensity when accomplishing an activity or solving a task (Figure 9). In essence, following this model, task performance has a curvilinear relationship to the electrodermal arousal values of participants and, respectively, to their stress and motivational level. It explains that people do their best performance when they experience the optimal level of mental pressure; this optimal zone helps people in maintaining motivation, engagement, and happiness at an ideal level. Conversely, if the pressure is too low, people are bored and exhibit lowered performance; if they experience too much pressure, they feel highly distressed, anxious, and unhappy, and consequently their performance decreases. Respectively, if their motivation level is not too low or too high (i.e., when assigned goal-free tasks), users show their best performance.

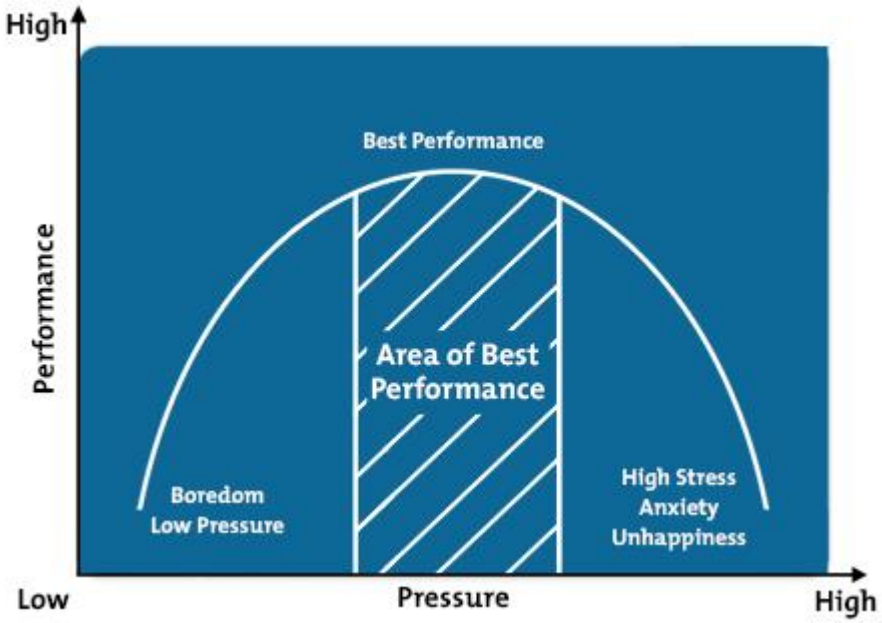


Figure 9: *The Inverted-U Model (Yerkes, Dodson (1908), as modified by Mind Tools https://www.mindtools.com/pages/article/inverted-u.htm, last access: 18.01.2017).*

Task performance might be influenced by other factors, such as the training level of the user and the difficulty of the task. According to Csikszentmihalyi's *Flow Theory* (1990), optimal performance occurs when the difficulty level of the task and the training level of the participant are in balance, as shown in Figure 10. Good balance exhibits flow when users are completely immersed in the task and show focused motivation. In other words, flow corresponds to positive excited arousal, in which the engagement level of participants (i.e., positive valence of their affective state) is high, and distress or boredom (i.e., negative valence) is low. When the difficulty or training level of participants increases, then their affective state might switch between boredom, engagement, and distress. For example, if highly or minimally trained participants switch from an easy task to a difficult one, then their arousal might increase, and the valence of their affective state might change from positive (i.e., engagement) to negative (i.e., distress). Conversely, by engaging in an easy task and by increasing the training level, participants might switch from an engagement state to boredom. On the one hand, task performance might be influenced by task difficulty and training of the participants, as well as by the valence of their affective state. On the other hand, the affective state influences task performance.

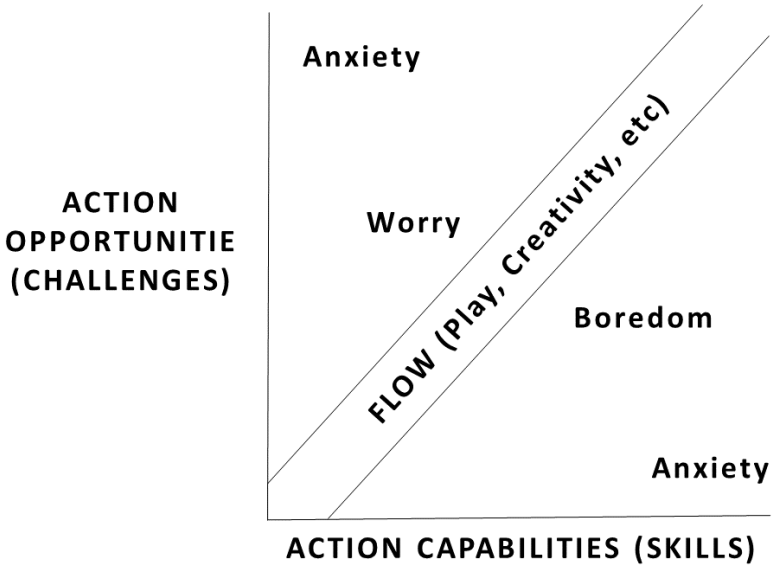


Figure 10: *Arousals versus difficulty level and participant skills (from Csikszentmihalyi 2000).*

Boyle et al. (2015) and Matthews et al. (2013) argue that engagement and distress are correlated with task performance. However, the consequences that worry might have on cognitive processes is not well understood to date (Barron et al. 2011). As Matthews et al. (2013) suggest, task engagement is correlated with energetic arousals, motivation, and concentration, and is often associated with attentiveness, joviality, and low fatigue. Distress is indicated by tense arousals, low hedonic tone, and low confidence-control, and it is often associated with low serenity and anxiety. Fairclough and Venables (2006) demonstrated that most of the physiological EDA measures (53%) are in relationship with engagement, and the rest (42%) with distress. Therefore, a positive affective valence corresponds approximately to a high level of engagement and a low level of distress, whereas a negative affective valence corresponds to a low level of engagement and a high level of distress.

According to Russell's *Circumplex Model of Affect* (Russell 1980) and Larsen and Diener's *Compromise Circumplex* (Larsen & Diener 1992), human emotions can be placed into a two-dimensional space having as axes bipolar valence (ranging from negative to positive) and physiological arousal or electrodermal activity (EDA) (ranging from low to high). For example, as seen in Figures 11 and 12, people presenting an affective state characterized by high arousals and positive valence means that they might be happy or satisfied. Conversely, people presenting an affective state characterized by low arousals and negative valence means that they might be bored or tired. In addition, simultaneously having high arousals and negative valence might correspond to a distressed state.

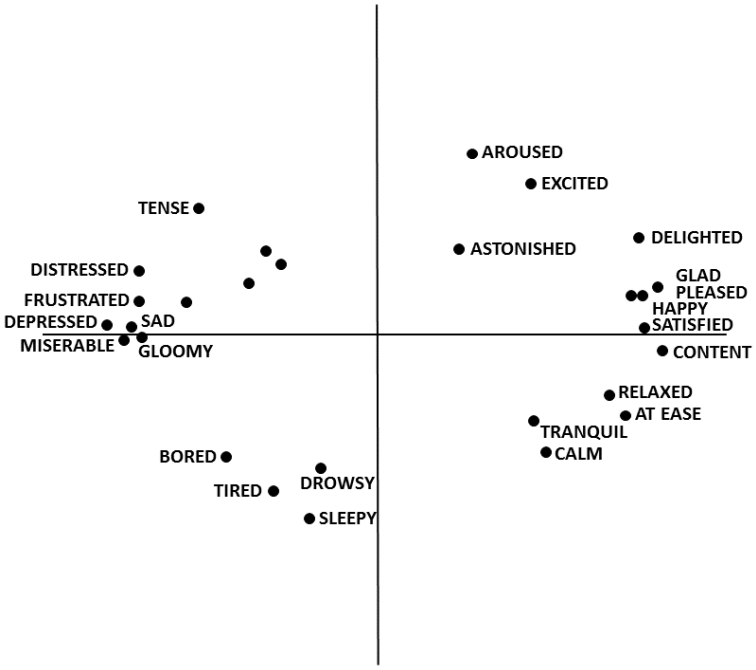


Figure 11: *Russell's Circumplex Model of Affect as a function of pleasure-displeasure (horizontal axis) and degree of arousal (vertical axis) (from Russell 1980).*

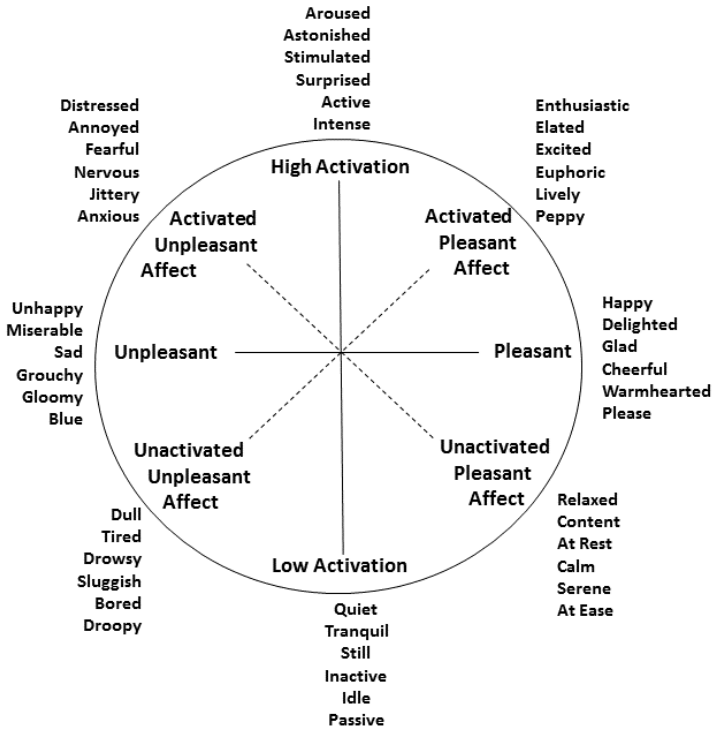


Figure 12: *Larsen and Diener's Compromise Circumplex (from Larsen & Diener 1992).*

To summarize, the following research gaps, which have been used to generate our research questions (cf. 1.5 "General Research Questions"), arose from the literature review presented above:

(1) **Use context, task and movement data:** Identify *for what type of use context, tasks and movement pattern* are animations particularly useful to effectively make visuospatial decisions.

(2) **Animation display design:** Assess what are the most *appropriate animated transitions* (or rate of changes) and *depiction of path history* for effective visuospatial decision-making with animations, especially of real-time spatio-temporal data. Create empirically validated design guidelines for perceptually salient, affectively engaging, and cognitively inspired animations.

(3) **User characteristics:** Better understand *how animation design is interconnected with perceptual, cognitive, and affective states of users* in a specific user context to support them in effective decision-making with animations. In particular, find out how animations might be effectively designed to *promote user motivation and engagement*.

In the next chapter, the requirements analysis and use case for the user studies of this thesis are described.

CHAPTER THREE

REQUIREMENTS ANALYSES AND THE AIR TRAFFIC CONTROL USE CASE

In this chapter, the results of semi-structured interviews with experts as well as use cases in the domain of Air Traffic Control (ATC) chosen for our two user studies of this thesis are presented.

3.1 Semi-Structured Interviews with ATC Experts

To effectively design the user studies proposed in this thesis, and thus increase their validity, different potential use cases have been analysed and different experts have been interviewed using semi-structured interviews (Barkley 1991). The experts that we contacted are researchers or operators in the domain of sport analysis, transportation systems and ATC. All of them are involved in decision-making with animations and real-time movement data. The purpose of these interviews was to identify potential needs, design issues, and realistic tasks to develop ecologically validated experiments with animations.

Because of many relevant affinities with our purposes and needs, we chose to only collaborate with ATC experts in the development of an empirical framework for this thesis. We thus conversed with two ATC researchers and one ATC operator about the importance of controller expertise, daily tasks, and animations for the effective monitoring of air traffic movement patterns and the prompt identification of air traffic conflicts. First, the interviewees explained to us that air traffic controllers use *semi-static animated displays as a standard display type* to monitor the air traffic space. However, the interviewed experts hypothesized that continuous animations (not used at the operational level yet) might be better than semi-static animations to successfully visualize aircraft movements and effectively detect anomalous or critical movement patterns. The continuous representation of movement dynamics might allow air traffic controllers to more quickly and more accurately detect speed and direction changes, as well as to better conceive of relative motion between aircraft.

According to the interviewed experts, we further identified two specific ATC tasks that are particularly relevant in their everyday job:

(1) The *prompt detection of aircraft movement changes* (speed and direction changes), and

(2) The prompt identification of critical situations in the air traffic flow by constantly *monitoring distances between aircraft* and by *predicting future aircraft movement patterns*, in particular, ensuring that aircraft maintain a minimum separation distance.

The correct perception of the relative motion between aircraft and of movement changes (e.g., speed changes) are thus critical tasks for ATC experts at the operational level. Regarding movement changes, as highlighted from the interviewed

ATC expert, *aircraft speed changes* are especially important for air traffic controllers to effectively and efficiently detect to have an accurate situational awareness (SA) and, thus, to successfully make critical decisions. For example, making adequate speed change manoeuvres for the resolution of aircraft conflicts is important for safety and economic efficiency (i.e., extra time and fuel consumption) reasons (Cetek 2009). This relevance was also confirmed in SA research (Endsley 1995) and in the post-test questionnaires with ATC experts after our experiments, as shown in the next section.

3.2 Use Case: Tasks, Movement Data, Animation Displays and User Characteristics in ATC

As mentioned in the previous section, we thus identified ATC as a real-world use case to test animations of movement data for our two user studies. We also identified two real-world SA tasks that are particularly appropriate for our purposes. These two tasks are examined in more detail below.

3.2.1 Tasks and Movement Data

As air traffic controllers are exposed to high-stress conditions, high cognitive workload, and time-constrained decision-making situations, an important study topic in ATC research is the level of SA that can be maintained by decision-making tasks. SA can be defined as 'the result of the continuous extraction of environmental information, integration of this information with previous knowledge to form a coherent mental picture, and the use of that picture in directing further perception and anticipating future events' (Dominguez et al. 1994). According to Endsley (1995), SA can then be divided according to three main states or levels: (1) *perception*, (2) *comprehension*, and (3) *projection into the near future* of the resulting spatio-temporal event under study. These levels can be achieved simultaneously, and the higher levels cannot be achieved without having first completed the previous level, as they are partly dependent on each other. A complete SA requires the accomplishment of all three of these levels, i.e., users might project future states of a spatio-temporal phenomenon only by having effectively perceived and understood the current situation first (Grier 2015).

According to the expert interviewees, the two main tasks of air traffic controllers are the prompt identification of anomalous aircraft movement patterns and to assure that

aircraft maintain a minimum safe distance between each other. These two tasks correspond to the first and third SA levels, and are used as real-world tasks for our two user studies.

Concerning air traffic movement data displayed and processed on ATC displays, we identified the following movement parameters and movement pattern types relevant for our user studies, and according to the two SA tasks mentioned above and the Dodge et al.'s *Taxonomy of Movement* (2008):

(1) (The accurate apprehension of) *aircraft speed and heading (i.e., velocity) changes of individual aircraft trajectories* (in accord with the first SA task).

(2) (The accurate apprehension of) *relative motion* between aircraft trajectories (i.e., *converging movement pattern*). This is also the basis for the third SA task, and concerns the effective prediction of future aircraft movement patterns, and thus the rapid identification of potential air traffic conflicts (Higgins & Ibrahim 2014).

SA further highlights the importance of graphic displays in allowing air traffic controllers to build an accurate mental representation of the current air traffic situation, but also to effectively predict future spatio-temporal trends of air traffic flows (Lowe 2015). As mentioned in Section 2.1, *Factors related to the Use, Task and Movement Data Context*, animations of movement data are particularly suited to *uncover speed patterns of single moving objects*, and to *recognize converging movement patterns of a group of moving entities* (Gudmundsson et al. 2012). These two actions, in turn, correspond again to the two SA tasks chosen for our user studies. The next section explains in more detail the animation design of standard semi-static ATC radar displays, and the potential of novel continuous animated displays.

3.2.2 **Animated ATC Radar Displays**

Standard ATC radar displays used to monitor real-time air traffic dynamics typically depict information by means of *semi-static animations*. Due to technical reasons, current aircraft positions are refreshed every 4 or 8 seconds, like a sequence of static images. In this case, the dynamic variable 'rate of change' corresponds thus to an abrupt transition between successive display scenes, which occurs every 4 or 8 seconds. In addition, past movements are depicted by means of symbols showing the last 5–8 aircraft positions (*path history*), and future positions can be interpolated by computational estimations of the current and past dynamics by means of a speed vector. For example, as shown in Figure 13, French ATC radar displays show aircraft dynamics according to the *Operational Display System* (ODS)

comet design (Hurter & Conversy 2008). The ODS comet radar displays depict each aircraft position with five squares of different size. The largest square corresponds to the current aircraft position, whereas the successive squares of gradually smaller sizes correspond to the aircraft's past positions (i.e., path history). Close to the depicted current and past positions is also indicated additional information, such as the aircraft's relative current speed and height, and aircraft type. A speed vector is also depicted by means of a line extended from the aircraft's current position to the desired future position. Usually, the operator can choose between a one, and a three-minute interval from the current status of the predicted future position.

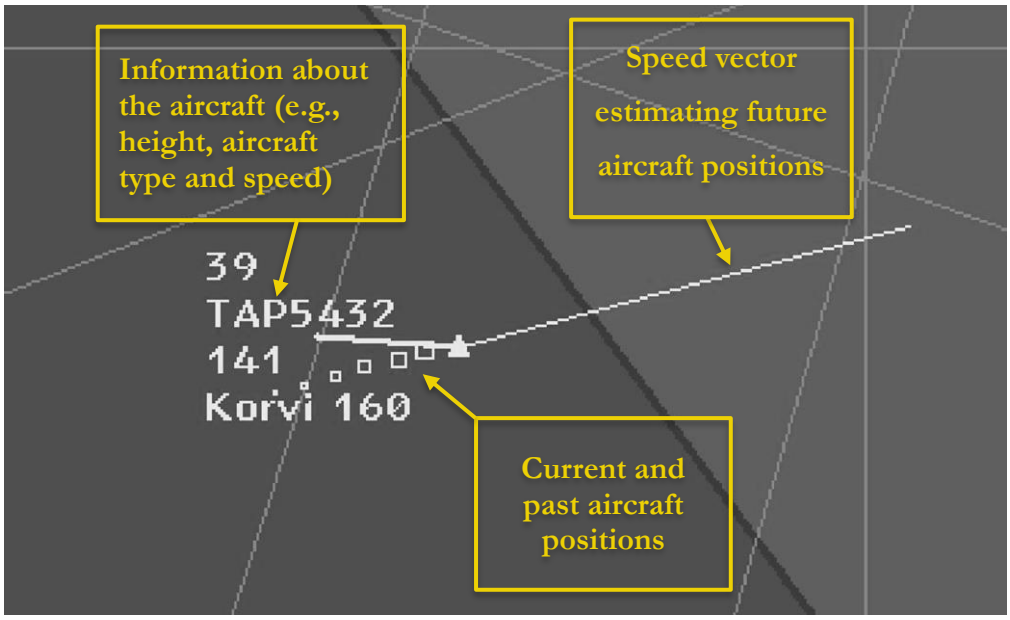


Figure 13: *An aircraft as represented on a typical French radar screen, where the largest square corresponds to the current aircraft position, the following four smaller squares represent past positions, and the speed vector predicts future movements (after Maggi et al. 2016). Annotations shown in yellow are for explanation only and do not belong to the display.*

Continuous animations of aircraft movements exist, such as the Flightradar24 system[1], but they are not currently used at the ATC operational level. What if continuous animations were used instead of semi-static animation at the ATC operational level? Lee and Klippel (2005) argue that continuous motion of the spatio-temporal information might be advantageous for air traffic controllers compared with standard semi-static displays in effectively and efficiently perceiving aircraft dynamics, because relevant spatio-temporal

[1] Flightradar24 by: https://www.flightradar24.com/.

information is refreshed frequently and it might help them to create a better mental picture of the air traffic dynamics. In addition, Schlienger et al. (2007) point out that continuous animations might improve the perceptual accuracy of aircraft movement changes. However, the above-mentioned authors state that there have been few empirical studies conducted about the second and third SA levels, i.e., whether animations are effective at supporting users in the comprehension and meaning of event changes, and in the projection of event states in the near future. Furthermore, past studies analysed continuous animations almost exclusively with abstract map scenarios instead of realistic ones, and lacked realistic tasks performed in a specific context of use.

3.2.3 Characteristics of Air Traffic Controllers

It requires a quite long training phase before becoming an air traffic controller. In addition, air traffic control candidates have to pass specific spatial ability tests (e.g., tests of mental rotation and visuo-perceptual speed abilities) to enter ATC training schools. Once operators, they have to continuously process the current spatio-temporal information displayed on ATC radar screens and make critical decisions. It thus presupposes a high cognitive workload, because of the high amount of information to process simultaneously, and the high responsibility required of the operators to assure air traffic security. For this reason, air traffic controllers often possess superior spatial skills, and a high resistance to stressful situations.

We believe that the presented use case in the ATC domain is a valid one for our empirical work, which is aimed at assessing animations of movement data for effective decision-making. By choosing real-world tasks, real-world animated displays and expert viewers allows us to test viewers in a familiar use context, and thus to perform an *ecologically valid research approach*. The next chapter illustrates the methodology of our empirical approach in more detail.

CHAPTER FOUR

METHODOLOGY

In this chapter, I introduce the methodology adopted to assess animated displays and their relationships to the two VA dimensions, 'user characteristics' and 'use context and tasks", according to the research questions postulated in Section 1.5, *General Research Questions*.

Traditionally, animations have been empirically assessed by means of quantitative research methods. Recently, geo-visualization researchers have employed more qualitative analysis (e.g., interviews and questionnaires) or mixed research designs combining both quantitative and qualitative methods (Knigge & Cope 2006; Štěrba et al. 2014; Çöltekin 2015). Mixed approaches may take advantage of both methodologies by providing researchers with more information about the phenomena under study (Çöltekin 2015). According to Rohrer (2014), qualitative methods may help researchers in better understanding underlying reasons and motivations, whereas quantitative methods might provide researchers with useful information to answer 'how much' and 'how long' types of questions. The use context and task type chosen are also relevant to effectively assess a specific animation design (cf. Section 2.1 "*Factors related to the Use, Task and Movement Data Context*"). Recently, different researchers (Davies et al. 2014; Fish 2015; Risko & Kingstone

2015) proposed to empirically test geo-visualization displays in more realistic use contexts and with real-world tasks. Moreover, the assessment of animations requires new empirical approaches to help cartographers better understand perceptual, affective, and cognitive processes of users during visuospatial decision-making with animations (Maggi et al. 2014, 2016).

Most of the empirical studies that have been conducted to date are based on the evaluation of the *general abilities* of users (i.e., their perceptual and cognitive skills) by solving a specific visuospatial task with maps. Perceptual and cognitive skills are mainly measured with respect to user task performances, i.e., according to the effectiveness (i.e., response accuracy) and efficiency (i.e., response time) of participant responses (Garlandini & Fabrikant 2009). However, task performance measures can also be coupled with eye movement and psycho-physiological metrics (e.g., recording of brain and electrodermal activity, and self-reported questionnaire answers) to gain a better understanding of the perceptual, affective, and cognitive processes of users, as well as their visual strategies (Duchowski 2007; Çöltekin et al. 2010; Holmqvist 2011; Maggi & Fabrikant 2014b). The implications of affective states (e.g., motivation, stress and engagement) of users on animation designs, and vice versa, has not been systematically assessed in GIScience and cartography to date (Fabrikant et al. 2012).

For these reasons, for our empirical work, we developed a holistic approach that combines both quantitative and qualitative methods, based on a novel triangulation analysis that couples different empirical data sources (i.e., psycho-physiological measurements and eye-tracking data) with traditional paper-and-pencil questionnaires (Maggi & Fabrikant 2014b). We designed two user studies in the context of ATC to examine how the three VA dimensions, 'use context/task', 'animation display design' and 'user characteristics may influence the *efficiency* and *effectiveness* of visuospatial decision-making with animated displays. Table 1 gives an overview of the tested elements of each VA dimension.

Table 1: *Overview of the relevant factors of the three VA dimensions, 'use context/task', 'animation display design' and 'user characteristics' tested in our two user studies.*

VA dimensions		
Context of use and tasks	**Animation display design**	**User characteristics**
User domain: ▪ *ATC.*	**Dynamic visual variables**: ▪ *Rate of change* of a display scene (or the smoothness of transitions between scenes): semi-static vs continuous animations.	**Individual differences**: ▪ User *spatial abilities*. ▪ User *affective state*: engagement, distress, and worry. ▪ User *cognitive state*: cognitive workload and motivation.
User tasks: ▪ Two SA tasks relevant for ATC: *apprehension of movement changes*, and *prediction of future movement patterns*.	**Static visual variables:** ▪ The depiction of the moving object trace (or *path history*).	**Group differences**: ▪ *User expertise* (or also training and familiarity): experts vs novices.
Conceptual modelling of the movement data and movement space: ▪ Data or moving object (MO) type/number under study: *two, four and eight moving aircraft*. ▪ Movement pattern type: *aircraft speed changes* (or accelerations), and *relative motion* of two converging aircraft.		

In the next section, I present first the adopted quantitative approach, and successively qualitative methods, employed for the two experiments. In the last section of this chapter, I describe cross-validation approaches that we used to combine qualitative and quantitative methods.

4.1 Quantitative Approach

In the *first experiment*, we examined how animation design (i.e., semi-static vs. continuous animation), task difficulty level (i.e., 4 vs. 8 displayed objects and relative motion of aircraft), and user-related factors (i.e., individual and group differences) influenced the efficiency and effectiveness of movement change apprehension with animations. In the *second experiment*, we tested effectiveness of users in predicting future movement dynamics of two converging aircraft with animated displays. Table 2 summarizes the SA tasks, the independent variables, the test stimuli, and the dependent variables of the two user studies.

Table 2: *Overview of the experimental variables for the two experiments.*[2]

	SA Tasks in ATC	Independent variables	Test stimuli	Dependent variables
Experiment I	*Apprehension of aircraft movement changes*	▪ 2 animation design types (i.e., semi-static vs. continuous animations). ▪ 2x2 task difficulty levels (i.e., 4 vs. 8 aircraft; same vs different relative speed). ▪ 2 ATC expertise levels (i.e., ATC experts vs. ATC novices).	▪ *16 animations*, of which: ▪ 16 semi-static and 16 continuous animations (between-subject design). ▪ 8 animations depicting 4 aircraft, and 8 animations depicting 8 aircraft (within-subject design). ▪ 4 animations depicting aircraft moving at the same speed, and 12 animations depicting aircraft moving at different speeds (within-subject design).	▪ Response accuracy ▪ Response time ▪ Eye movements ▪ EDA ▪ EEG ▪ SSSQ ▪ Spatial abilities ▪ Self-reported metrics
Experiment II	*Prediction of future aircraft movements*	▪ 2x2 animation design types (i.e., semi-static vs. continuous animations; with vs. without path history). ▪ 2 task difficulty levels (i.e., same vs. different relative speed; 3 minimum separation distances). ▪ Only ATC experts tested.	▪ *24 animations* (within-subject design), of which: ▪ 12 semi-static animations, and 12 continuous animations. ▪ 12 with and 12 without path history. ▪ 12 animations depicting aircraft moving at the same speed, and 12 animations depicting aircraft moving at different speeds ▪ 8 animations for each of the three tested minimum separation distances.	▪ Response accuracy ▪ Eye movements ▪ EDA ▪ SSSQ ▪ Self-reported metrics

[2] EDA = Electrodermal activity.
EEG = Electroencephalogram or encephalography.
SSSQ = Short Stress State Questionnaire (Helton 2004).

4.1.1 **Measuring Eye Movements**

Standard eye movement metrics are more suitable for static displays. Eye movement analysis on animated displays is quite challenging, because data are more complex than those on static displays (Maggi et al. 2016). For example, standard metrics usually do not consider eye movement patterns over time, but only their spatial distribution (collapsed over time). However, the temporal relationships combined with spatial information among fixated objects might be very useful to understand how and when participants make decisions with the tested displays. For this reason, new methodologies are needed to analyse eye movements on animations.

We recorded user eye movements with the *Tobii Eye Tracker TX300*.[3] We analysed the collected data with *Tobii Studio 3.2.2 software*.[4] From these data, different standard metrics can be extracted to analyse the spatial distribution of eye fixation on the tested displays, including fixation count, fixation duration, and fixation rate. Further, it is also possible to examine eye movement sequences as indicators of strategies employed by the participants in visuospatial tasks, e.g., by means of *transition matrices* (Ponsoda et al. 1995; Çöltekin et al. 2010). For our eye movement data analysis, we further used a new methodology developed by Krejtz et al. (2014) that uses *entropy metrics* to compare individual eye movement across participants. These metrics have the advantage to more deeply investigate the sequences of participants' eye movements, compared to standard eye movement analyses, and, consequently, to better understand the strategies used by participants to solve the required tasks.

---

[3] Tobii TX300 by: http://www.tobiipro.com/product-listing/tobii-pro-tx300/.

[4] Tobii Studio by: http://www.tobiipro.com/product-listing/tobii-pro-studio/.

4.1.2 **Measuring Electrodermal Activity**

For the first time in cartography research, we recorded the electrodermal activity of participants during both experiments. Our purpose was to measure the skin conductance responses of participants during information processing with animations as an indicator of their cognitive load and their emotional reactions (Holmqvist 2011). Electrodermal activity was recorded with 1) a *Smartband* (in Experiment I) (Papastefanou 2009),[5], and 2) with a *E4* device (in Experiment II),[6] both of which are wearable sensor-wristbands.

Both bands are provided with electrodes and thermal sensors, by which various skin proprieties and physiological parameters can be measured, e.g., the users' electrodermal activity (EDA). EDA is defined as 'autonomic changes in the electrical properties of the skin', or, most commonly, as '*skin conductance*' resulting from sympathetic neuronal activity (Braithwaite et al. 2013). Skin conductance is an indicator of changes in sympathetic arousal and thus of *emotional state* (e.g., anger, calmness, distress and engagement), that is, *emotional responses* that occur also without cognitive mechanisms (e.g., threat, anticipation, salience, and novelty), and of *attentional processing*, where high EDA levels show salient stimuli or difficult task levels (Braithwaite et al. 2013).

Skin conductance is usually measured as tonic EDA (*skin conductance level*, SCL) and phasic EDA (*skin conductance response*, SCR), as shown in Figure 14. An increase of SCL/SCR is an indicator of increased EDA intensity, and a decrease of SCL/SCR is an indicator of decreased EDA intensity. To systematically quantify the intensity of EDA responses, Figner and Murphy (2010) proposed calculating the integral of positive SCR values over a specific period of time (i.e., the *area under the curve*, or the sum of all momentary skin conductance change values, over the stimulus duration, AUC), as shown in Figure 15. AUC scores can also be linked with decision-making processes (Figner, Murphy 2010). I discuss more in detail about this possibility in Section 4.3, *Triangulation Approach*. Further, to compare individual electrodermal activity (EDA) across groups Lykken (1972) suggested normalizing SCR values as phi-scores by means of the range correction transformation procedure. The complete procedure to analyse EDA data is described step by step in Annex 8, *EDA Analysis*.

---

[5] Smartband by: http://bodymonitor.de/.

[6] E4 by: https://www.empatica.com/e4-wristband.

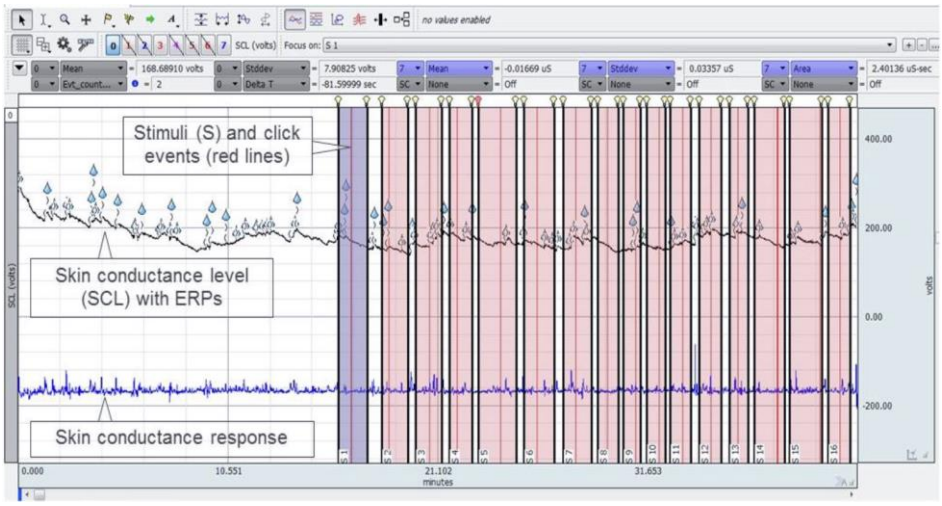


Figure 14: *Example screenshot of typical individual EDA responses (SCL and SCR) during the first user study with 16 animations (ERP = Event-Related Potential) analysed with AcqKnowledge 4 (from Biopac),* [7] *(from Maggi & Fabrikant 2014b).*

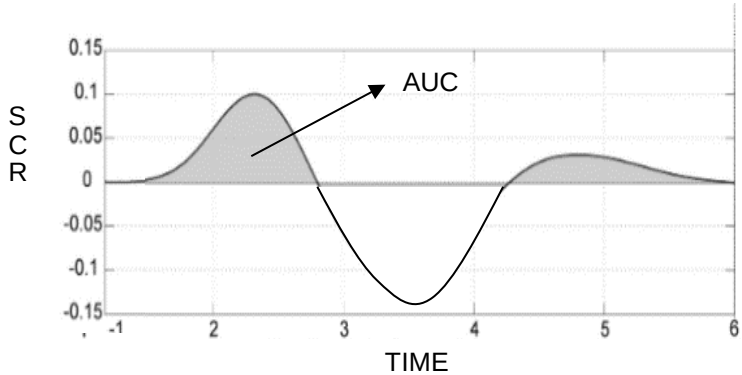


Figure 15: *Example of skin conductance responses (SCRs) measured over 6 seconds with positive AUC values in grey. The integral of the positive AUCs corresponds to the EDA intensity (after Figner & Murphy 2010).*

[7] AcqKnowledge by: https://www.biopac.com/product/acqknowledge-software/.

4.1.3 Measuring Brain Activity

EEG research has a long tradition of studying human brain activity related to cognitive and emotional processes, such as cognitive load and affective states (Benedek et al. 2014). EEG data are filtered by means of standard analysis procedures into different spectrum frequency bands (i.e., alpha, beta, gamma, and theta) (Delorme & Makeig 2004), and specific psycho-physiological processes can be revealed (Antonenko et al. 2010).

During the first experiment, we also measured participants' brain activity to determine their *cognitive load* and the *valence of their affective states*. We recorded EEG data with the mobile device Emotiv EPOC.[8] Cognitive load can be analysed by extracting alpha power values from raw EEG data (Benedek et al. 2014), while *approach-related motivation* can be measured by means of *Frontal Alpha Asymmetry* (FAA) measures (Briesemeister et al. 2013). The complete procedure to analyse EEG data is described step by step in *Annex 9: EEG Data Analysis*.

Furthermore, Gable and Harmon-Jones (2008) highlight the interconnection between FAA and memory performance, as well as between FAA and visual attention. They found that more positive, left-sided asymmetry scores might be indicative of increased memory and attentional performance, higher approach motivation/engagement level, and more focused task attention.

4.1.4 Measuring Users' Spatial Abilities

As described in Section 2.3, *Factors related to the User*, there are different ways to measure spatial ability, depending on the studied displays and tasks. In our first experiment, we measured participant visuo-perceptual speed by means of the Hidden Pattern Test (Ekstrom et al. 1976). A subset of this test is provided in Annex 2, and the complete test is available online at: http://www.ets.org/Media/Research/pdf/Kit_of_Factor-Referenced_Cognitive_Tests.pdf (last access: 18.01.2017).

[8] Emotiv EPOC by: https://emotiv.com/.

4.1.5 **Measuring Task Performance**

Usability metrics are often employed in cartography to compare different map design types and to assess participant task performances, such as how effectively and efficiently they solve the required task. The most useful metrics to compare multiple designs and tasks are *issues-based metrics* (Tullis & Albert 2008). According to Štěrba et al. (2014), three of the most widely employed usability parameters to measure the quality of cartographic products, engaged especially in usability engineering studies, are (Fish 2015):

- *Effectiveness* – Being able to complete a task, or the **task success**. Recommended question for assessment: 'I am satisfied with the ease of completing the tasks in this scenario' (Lewis 1990).
- *Efficiency* – The amount of **effort** required to complete the task (completion time, but also number of pages/objects views). With eye tracking metrics for cognitive load or ease of use: blinks, saccades, and fixations. Recommended question: 'I am satisfied with the amount of time it took to compete the task in this scenario' (Lewis 1990).
- *Satisfaction* – The degree to which the user was happy with his or her experience while performing the task.

In our two experiments, we assessed the first two usability parameters, participant effectiveness (i.e., response accuracy) and efficiency (i.e., response time, or task completion time), in solving real-world tasks with realistic ATC animations.

4.2 Qualitative Approach

Beyond the quantitative methods mentioned in the previous section to assess animations, our empirical approach was based on qualitative assessments using post-test questionnaires.

Through such questionnaires, it is possible to elicit motivations, opinions, preferences, attitudes, etc. from participants. According to Albert and Dixon (2003), the following three usability metrics are well suited to measure subjective design preferences and confidence:

- *Ease of use*: How easily participants solve the required task with a specific animation design type?
- *Usefulness*: How useful is a specific animation design type to solve the task?
- *Self-Confidence*: How is their confidence in using a particular animation design type and in solving a specific task?

For both of our user studies, we asked participants to complete a post-test questionnaire containing questions designed to assess the above-mentioned metrics. All of the usability test questions are shown in Annexes 4, *Post-Test Questionnaire of Experiment I*, and 5, *Introduction and Questionnaire for Testing Usability Metrics of Experiment II*.

Beyond general usability metrics, a more detailed questionnaire might also be created to better identify a broader personal experience for a user with a specific interface. User experience (UX) encompasses different user-related factors, e.g., expectations, motivations, perceptions, and emotions that users experience by using the tested display. User experience and usability are interconnected. If a specific interface allows users to perform the task more easily compared to another interface, then they might be more motivated to solve the task. Conversely, if an interface makes it more difficult to solve a task, a user might be frustrated and no longer motivated to solve it. This might lead to the following question: What is the most effective method to assess the UX during an experiment with specific animated displays?

One way to measure the UX of participants during an experiment, in particular their emotional and affective state, can be through standard self-reported questionnaires. For example, we measured the perceived stress-levels of participants by means of the *Short Stress State Questionnaire* (SSSQ) (Helton 2004). Before and after the experiment, participants were asked to answer 24 questions in which they reported, e.g., whether they felt active, annoyed, or motivated when doing the task. All of the SSSQ questions are shown in Annex 1.

Another way to measure the UX of participants during an experiment might to couple the abovementioned self-reported questionnaires with emotional and cognitive processes measured with psycho-physiological sensors (i.e., EDA data), with devices measuring brain activity (i.e., EEG data) and eye movement data (Holmqvist 2011; Maggi & Fabrikant 2014b). This cross-validation analysis is explained in more detail in the next section.

4.3 Triangulation Approach

The triangulation approach allows researchers to assess the cognitive processes of participants, as well as their emotional and affective states (e.g., motivation, cognitive load, frustration, stress), with enhanced confidence and reduced uncertainty of interpretation (Webb et al. 1966; Holmqvist 2011). As Holmqvist (2011) points out, eye tracking data, as a single data source to analyse users' cognitive mechanisms in HCI user studies, may be useful to determine *where* and *how long* a specific subject's cognitive process has been activated. However, it is not sufficient to tell us *which* process and *why* it is involved.

The recent emergence of psycho-physiological monitoring technologies allows the coupling of these technologies with eye tracking data, as well as with qualitative records form self-reported questionnaires. For example, EEG data might be coupled with EDA and eye tracking data to reliably assess cognitive workload, stress, fatigue, or engagement of test participants. In addition, by coupling these data with qualitative records from self-reported questionnaires, it is possible to perform a cross-validation analysis that can not only explain the *type of cognitive and emotional process* involved in a specific task, but also the reason (the *why*) for a good or bad task performance and a participant's design preferences (Maggi & Fabrikant 2014b).

The triangulation approach between eye movement data with EEG and/or EDA data may be useful to better assess decision-making processes of participants dealing with animations over a specific time period (e.g., from the animation starting point to when the participant makes a decision) (Maggi & Fabrikant 2014b). We found that the strategies of participants to solve a task with animated displays might change over time and that this is reflected in eye movement and EEG/EDA data. Decision-making processes, as well as other cognitive processes, can be measured with EEG and EDA data through the *event-related potential* (ERP) technique (Luck 2005). The temporal progression, as well as the combination of these data, might facilitate the assessment of how and when exactly a

decision has been made with animations, and thus achieve the better evaluation of the implication of a specific animation design choice for decision support with real-time spatio-temporal data.

To summarize this chapter, in this thesis we developed two user studies in the context of ATC. We collected data from various sources (i.e., response accuracies, eye movements, EDA, EEG and post-test questionnaires) and we analysed data according to a cross-validation method. The use of EEG and EDA sensors to analyse participants' perceptual, affective and cognitive processes involved in animation tasks has not been studied in the context of cartography and GIScience to date. We believe that this empirical approach allows us to better understand interconnection between animation design, use context/task types and user characteristics, as well as decision processes with animated displays. The experimental design and results of the two experiments are presented in the next two chapters.

CHAPTER FIVE

EXPERIMENT I: APPREHENSION OF MOVEMENT PATTERNS WITH ANIMATIONS[9]

In this chapter, the experimental design and the results of the first experiment with air traffic control animated displays are described.

[9] Some of the content of this chapter has been published previously in Maggi et al. 2016, especially Sections 5.3, *Experimental Design* and 5.4, *Results*.

5.1 Experiment Background and Goal

This first user study is aimed at better understanding how the three main VA dimensions (i.e., animation design, use context and task, and user characteristics) might affect the apprehension span of task-relevant movement features of animations.

As explained in Section 2.1, *Factors related to the Use, Task and Movement Data Context,* and Chapter 3, *Requirements Analysis and The Air Traffic Control Use Case*, we decided perform our user study in the domain of Air Traffic Control (ATC) to adequately respond to our main research questions (cf. Section 1.5 *General Research Questions*). One of the main tasks of air traffic controllers is to identify abnormal movement patterns of aircraft. This is related to the first statement of spatial analytics (SA), which aims at analysing how well users perceive task-relevant objects on visual graphics. In particular, the *effective and efficient detection of aircraft movement speed* is a particularly relevant task in ATC, because it improves the controller's situational awareness. In this first experiment, we thus investigated how users apprehend speed changes (i.e., accelerations) of aircraft with animations with respect to the first SA statement.

Typically, air traffic controllers view animated displays that refresh every 4 seconds (or also 6 or 8 seconds) to monitor air traffic in real time, that is, semi-static animations. One question that was raised from our interviews with ATC experts was whether when using unfamiliar, continuous animations, air traffic controllers could more accurately apprehend aircraft movement changes (i.e., detection of aircraft accelerations), due to the quickly refresh rate of the animated display, and thus more realistic depiction of aircraft movement dynamics.

Air traffic controllers usually possess superior spatial skills and work under very stressful conditions, not only because of the importance of their task, but also because of the high cognitive overload due to the large number of objects that they have to process simultaneously. With this first experiment, we wanted to find out if air traffic controllers react differently with novel kinds of animated displays when solving typical SA tasks, compared to standard radar displays. In other words, we were interested in analysing to what degree the visual analytics (VA) dimension 'animation design', in this case the dynamic visual variable *rate of change* of the display, might affect participants' perceptual, affective (i.e., stress, engagement, and worry), and cognitive (i.e., motivation and cognitive load) processes linked with their task performance. In addition, we would like to examine if the

familiarity with ATC displays and tasks, and the superior *spatial skills* of ATC experts might help them in successfully solving the proposed tasks with novel animation design types. To do that, we compared their task performance with this of ATC novices (i.e., people without any experience in ATC).

The specific research questions for this user study, and derived working hypotheses, are presented in more detail in the next section.

5.2 Specific Research Questions and Working Hypotheses

The specific research questions for this first experiment, related in turn with the general research questions mentioned in Section 1.5, *General Research Questions*, are the following:

- **SRQ 1 (CONTEXT OF USE and TASK):** How does task difficulty (i.e., *number of depicted moving objects* and *relative motion*) influence the apprehension of movement changes in animations?
- **SRQ 2 (ANIMATION DESIGN):** How does the *animation type* (i.e., semi-static vs. continuous animations) influence the apprehension of movement changes?
- **SRQ 3 (USER):** How do specific *user-related factors* (i.e., perceptual, cognitive, and affective processes, including individual and group differences) influence the apprehension of movement changes in animations?
  - **SRQ 3.1:** How does *expertise* (or also previous knowledge, familiarity, and training*)* of the user (i.e., *experts vs. novices*) influence the apprehension of movement changes in animations?
  - **SRQ 3.2:** How do the *spatial abilities* of the user (i.e., *low vs. high spatial abilities*) influence the apprehension of movement changes in animations?
  - **SRQ 3.3:** How do *perceptual, cognitive, and affective states* of the user influence the apprehension of movement changes in animations?

From these specific research questions, we derived their corresponding working hypotheses as follows:

- First, we expected that the *task difficulty* (i.e., number of depicted moving objects and differences in their relative motion) influences the apprehension of movement changes. According to Ware (2013), only a few objects, especially when moving at similar relative speed, can be effectively processed simultaneously. Thus, an increase of depicted entities, as well as a decrease in the relative speed difference between the moving

objects, might influence negatively the perception of movement changes with continuous animations (effect of change blindness).

WH 1: *Users can effectively apprehend movement changes if few elements that also move at similar relative speed are visualized simultaneously (Koffka 1935; Ware 2013). Increasing the number of moving objects displayed on the screen, as well as the difference of their relative speed, negatively influences the apprehension of movement changes.*

- Second, we expected that the *animation design* has an influence on the apprehension of movement changes. With semi-static animations, users have to reconstruct aircraft dynamics from a sequence of static images; aircrafts dynamics can be inferred using the relative distance of past positions. With continuous animations, dynamics are perceived directly, naturally, and immediately. For this reason, users might more efficiently perceive spatio-temporal data with continuous animations than with semi-static ones (they satisfy the *Apprehension Principle* of good graphics, with reference to Tversky et al. 2002). However, continuous animations imply higher cognitive costs, because users have to process the depicted information continuously (Lowe 1999).

WH 2 (ANIMATION DESIGN): *Users can more easily apprehend movement changes with continuous animations than with semi-static animations (Tversky et al. 2002).*

- Third, we expected that experts would perform the required SA tasks more effectively and efficiently than novices due to their training and practice in ATC. In addition, experts might perform worse on tests with continuous animations because of their training being limited to semi-static animations (mismatch with controllers' heuristics, see Lee & Klippel 2005) and not because continuous animations are less efficient in perceiving movement changes. The spatial abilities of users could also influence the efficacy of one design compared to another.

WH 3.1: *Experts will perform tasks more effectively and efficiently with both animation types than novices due to their previous training (Hegarty et al. 2003).*

WH 3.2: *Experts might perform worse on tests with continuous animations because of their training being limited to semi-static animations (mismatch with controllers' heuristics, see Lee*

& Klippel 2005). In general, users will perform tasks more effectively and efficiently with semi-static animations if they have high spatial abilities. Respectively, users will perform tasks more effectively and efficiently with the continuous design if they have low spatial abilities (Hegarty et al. 2003).

- Finally, since with continuous animations users must process data continuously, this might negatively affect emotional and affective states (e.g., stress levels) of the users, and consequently their task performance. However, with semi-static animations, users must wait 4 seconds to visualize the future position of an aircraft, and this might negatively affect users' affective state (e.g., stress increases). Conversely, a more realistic depiction of aircraft movements might thus positively affect the affective and cognitive state of the user, because it allows users in better form a mental representation of the current and future aircraft dynamics. Similarly, the pauses from one state to another might reduce the cognitive load and thus have a positive effect on the emotional and affective state of the user.

WH 3.3: *The type of animation with which users are asked to solve the required task, user expertise level and familiarity with the task might influence their emotional and affective states, and consequently their task performance.*

5.3 Experimental Design

We designed a mixed factorial experiment according to both a between- (for the independent variable *animation design*) and a within-subject design (for the other independent variables). The three independent variables (IV) that we tested are as follows:

- ***IV 1 (CONTEXT OF USE and TASK):*** We manipulated two factors: (F1) the number of the depicted aircraft, and (F2) the relative motion between moving aircraft (i.e., same vs. different speeds).

  The first factor, i.e., the **number of the depicted aircraft (F1)**, has two levels of difficulty:
  - **Difficulty level 1 (4O):** Four moving aircraft are depicted on the screen.
  - **Difficulty level 2 (8O):** Eight moving aircraft are depicted on the screen.

  The second factor, i.e., the **relative motion between moving aircraft (F2)**, has two levels of difficulty:
  - **Difficulty level 1 (ES):** All the depicted aircraft move at the same speed on the screen.

- **Difficulty level 2 (DS):** All the depicted aircraft move at different speeds on the screen.

- ***IV 2 (ANIMATION DESIGN):*** We manipulated the dynamic visual variable "rate of change" (or the smoothness of the transitions between scenes) of the animated displays.
  The first factor, i.e., the **rate of change of the animated displays**, has two levels:
  - **Level 1 (S):** Semi-static animations (standard animation design type that air traffic controllers use in their everyday jobs) (frame rate: 1/4 Hz).
  - **Level 2 (C):** Continuous animations (novel animation design type for air traffic controllers) (frame rate: 60 Hz).

The choice of the above mentioned independent variables and the development of the experimental design is the result of several discussions with researchers and air traffic controllers at the 'Ecole Nationale de l'Aviation Civile' (ENAC) in Toulouse (France), and as result of a pilot test with Swiss air traffic controllers ran at the University of Zurich. To summarize, Table 4 illustrates the independent variables used for this experiment, including their corresponding factors and levels.

Table 3: *Summary of the independent variables for the first experiment with their corresponding factors and levels.*

INDEPENDENT VARIABLE	FACTOR	LEVEL
(CONTEXT OF USE and) TASK DIFFICULTY	*Number of the depicted moving aircraft*	Aircraft moving at equal speed
		Aircraft moving at different speeds
	Relative speed	Aircraft moving at equal speed
		Aircraft moving at different speeds
ANIMATION DESIGN	*Rate of change (or smoothness of the transitions between scenes)*	Semi-static animations (1 display every 4 seconds); transitions between scenes are abrupt
		Continuous animations (60 displays every second; 60 Hz); transitions between scenes are smooth and continuous

5.3.1 **Participants**

A total of 37 participants participated in this first experiment, 18 experts in the ATC domain, and 19 novices. The experts were all trained air traffic controllers at ENAC in Toulouse (France), each with more than 10 years of experience and training in the context of air traffic control. On average, expert participants were 38 years old. Sixteen of the experts were male and two were female. Novices were undergraduate students at Temple University in Philadelphia (USA) without any prior knowledge of ATC. Novices were, on average, 22 years old. Novices were distributed equally across gender.

5.3.2 **Materials**

We created 16 animations for the study. We designed the animations according to the *Operational Display System* (ODS) comet design, currently used for air traffic control radar displays in France (cf. Chapter 3, *Requirements Analysis and The Air Traffic Control Use Case*). The test stimuli were created using Processing,[10] Java-based software that is particularly suited for developing animations and dynamic graphical interfaces.

As mentioned in the previous section, we manipulated the independent variables *animation design* by changing the dynamic visual variable 'rate of change' (or smoothness of the transitions between scenes) of the animated displays. In all, 16 semi-static animations and, respectively, 16 continuous animations were been created. In the semi-static animations, aircraft positions refreshed abruptly at 4-second intervals (i.e., 1 frame each 4 seconds), and in the continuous animations, aircraft positions refresh smoothly and continuously (i.e., 60 frames per second). This independent variable was tested according to a between-subject design, in which one group of participants (equally distributed with respect to experts and novices) ran the experiment with only the semi-static animations, whereas a second group of participants (equally distributed with respect to experts and novices) ran the experiment with only the continuous animations.

Further, we manipulated the independent variable *task difficulty* by changing the number of the depicted aircraft moving on the scene and their relative speeds. Eight of the 16 animations depicted four aircraft and eight displays visualizing eight aircraft, as shown in

[10] Processing at: https://www.processing.org/.

Figure 16. All aircraft moved with typical aircraft take-off speeds (i.e., 160 kts,[11] 200 kts, 250 kts, and 290 kts). In four animations, all aircraft moved at equal speed (ES condition), and in 12 displays, aircraft moved at different speeds (DS condition). In all animations, aircraft moved at constant speed. One aircraft of the tested stimuli started slowly, after four seconds, to accelerate. This accelerating aircraft was the thematically relevant object that participants are asked to identify. Accelerations were computed according to averaged real values for most common aircraft types (i.e., A320), 0.4 kts/s. Accelerations (or velocity movement changes) are interpolated from the starting position and the ending position of the aircraft during the animation.

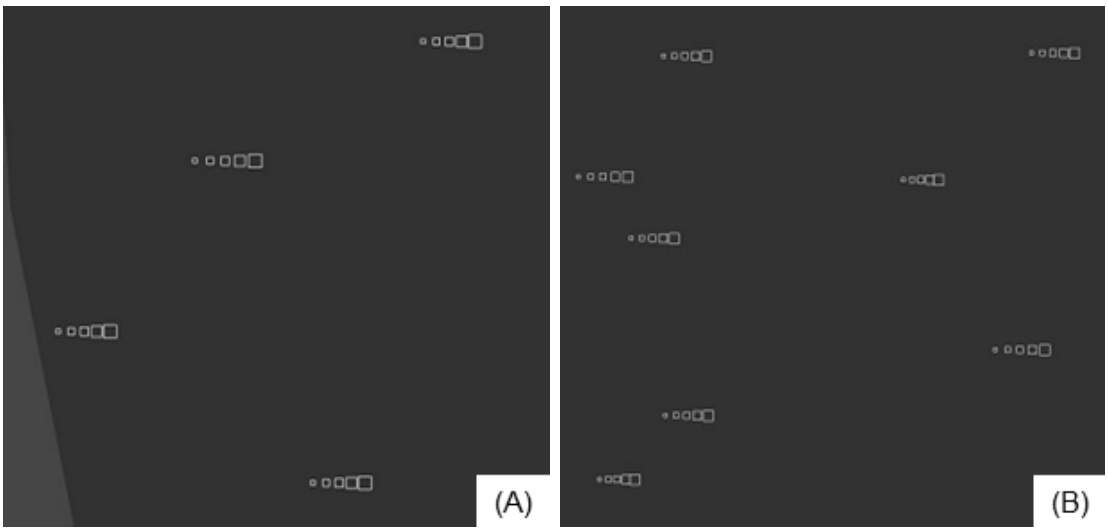


Figure 16: *Static representation of two stimuli with (A) 4 and (B) 8 aircraft moving from left to right on the screen at different speeds (from Maggi et al. 2016).*

To analyse a potential shift of attention, i.e., users focussing their attention mostly on perceptual salient objects (the fastest moving aircraft) rather than on thematically relevant objects (the accelerating aircraft), we designed the stimuli so that the accelerating aircraft was never the fastest aircraft. To assess the perceptual saliency of the animated scenes, we calculated dynamic saliency maps for each stimulus according to Itti et al. (1998), using the *iLab Neuromorphic Vision C++ toolkit.*[12] This software models human vision from a neuromorphic point of view. Figure 17 shows an example of a test stimulus with four aircraft moving at different speeds.

---

[11] Kts = knots (nautical miles per hour).

[12] iLab Neuromorphic Vision C++ toolkit: http://ilab.usc.edu/toolkit/.

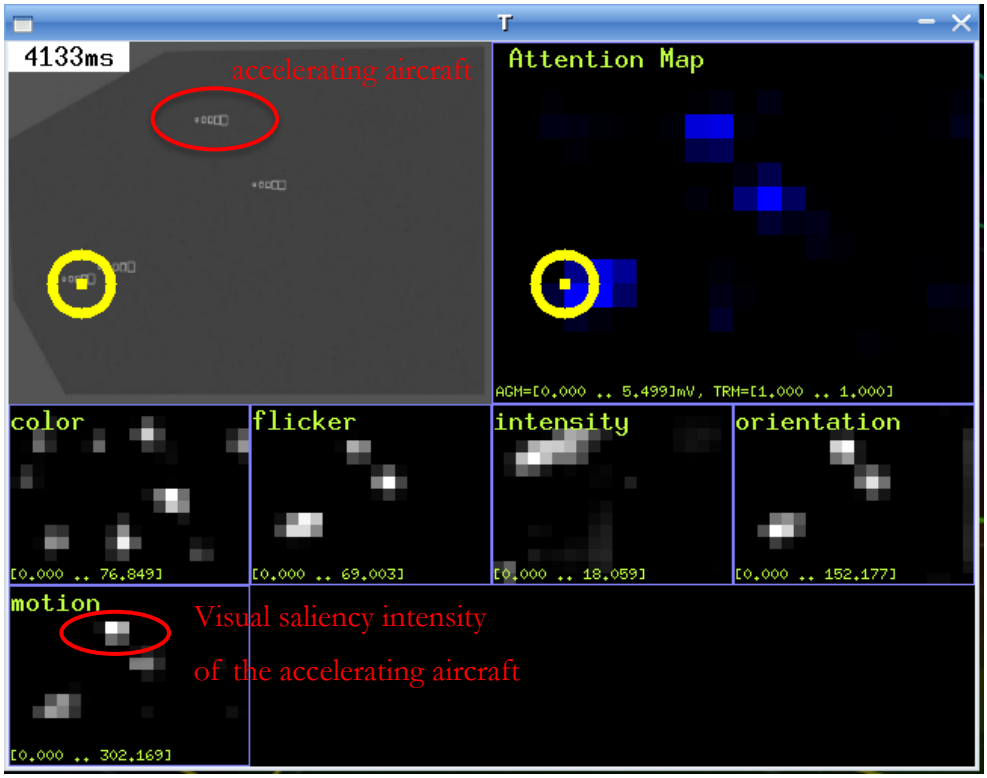


Figure 17: *Screenshot of an attention map of a test stimulus with four aircrafts moving at different speeds computed with the iLab Neuromorphic Vision C++ toolkit11. With this software it is possible to analyse the saliency of a moving feature (on the box "motion"). In this case, the accelerating aircraft should be perceived by users as the most salient moving aircraft.*

As shown in Figure 18, aircraft speeds and accelerations can be inferred visually from the total length of the ODS radar comet (D in Figure 18) and the spaces (d1 and d2) between its current and past positions (t1 and t2, respectively). A faster-moving aircraft has greater spacing between the 5 position squares, i.e., the spacing between d1 and d2 in Figure 18 is greater than it would be for a slower aircraft. Faster-moving aircraft thus also show a longer overall comet length (i.e., D in Figure 18) compared to slower-moving aircraft. Accelerating aircraft show a constantly increasing spacing between the 5 position squares; decelerating aircraft show the opposite (e.g., d1 is smaller than d2 in Figure 18). Accordingly, the total length of the comet (i.e., D in Figure 18) increases or decreases (Hurter & Conversy 2008).

With the two animation design types, i.e., semi-static and continuous animations, the perception of aircraft speed and acceleration is different. With semi-static animations, aircraft speeds and accelerations are perceived indirectly from the aircraft dynamics by inferring changes between the depicted position squares. In contrast, with continuous animations, aircraft speeds and accelerations are also perceived, in addition to the visual

cues provided by spacing between the position squares and the total radar comet length, directly from the aircraft dynamics. In other words, with continuous animations, speeds and accelerations are encoded twice, i.e., with the aircraft motion and graphically with aircraft position spacing and overall comet length.

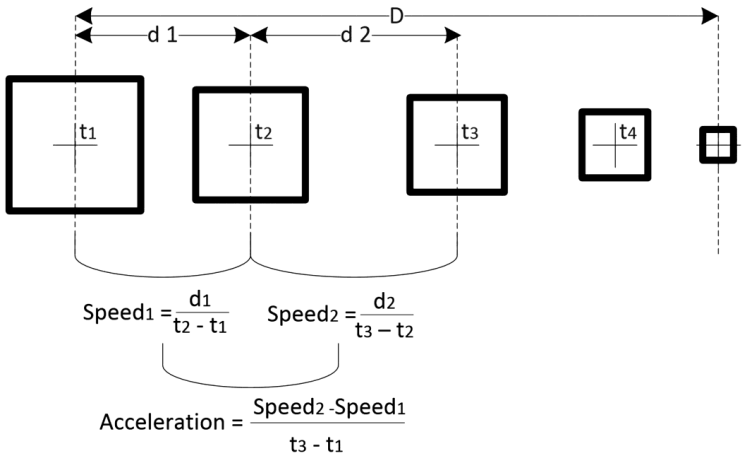


Figure 18: *Speed and acceleration of an aircraft as depicted on the French ATC radar display (ODS) (*D *indicates the overall comet length,* d *indicates the spacing between aircraft positions, and* t *indicates the temporal progression of the depicted aircraft positions) (after Hurter et al. 2008; Maggi et al. 2016).*

To analyse only the effect of the above-mentioned independent variables, all other potentially confounding variables are kept constant. Aircraft move all in the same direction, and they are depicted with the same colour as in the displays at ENAC (i.e., white aircraft moving on a homogeneous dark grey background). We also omitted text labels and speed vectors in order to avoid possible distractions.

5.3.3 Collected Data and Test Setup

We collected various data (i.e., dependent variables) to evaluate participant task performance and decision-making strategies. We recorded the response accuracy and response time of the participants. Response accuracy refers to the correct detection of the thematically relevant object (against visually salient distractors). Response time was computed as the time that participants took to complete the task (i.e., detection of the accelerating aircraft).

Visual spatial skills can be assessed by means of a standard spatial ability test (i.e., the Hidden Patterns Test) (Ekstrom et al. 1976), and was used in the present research. This test is particularly appropriate to measure the speed and accuracy of visual searches in

dynamic scenes. It consists of identifying a specific figure that is hidden amongst other elements. In our study, participants had to process 200 different patterns, and mark whether a figure was visible or not. In all, they had six minutes to solve the complete test. The complete Hidden Pattern Test can be seen in Annex 2: *Hidden Pattern Test*.

Moreover, participants' electrodermal activity was recorded by means of a 'Smartband', a skin conductance sensor,[13] and participants brain activity was recorded with Emotiv, to provide an electroencephalogram.[14] We also collected participants' eye movement data with a Tobii TX300 eye tracker.[15]

5.3.4 **Procedure**

Before initiating the main experiment, participants were welcomed to the lab and they were asked to fill in a background questionnaire (see Annex 3: *Pre-Test Questionnaire*). Next, they were asked to solve a Hidden Patterns Test in which their spatial abilities were measured. After a brief training phase, participants were asked to process 16 animations according to a between-subject design, i.e., they watched either 16 semi-static or 16 continuous animations. The required task consisted in detecting, as soon as possible, the accelerating aircraft (by clicking on it with a right mouse click). Animations were presented to them digitally and in random order on a colour monitor at a 1920 x 1200 pixel spatial resolution. Once participants had identified the target aircraft and confirmed their choice, the animation stopped. Following that, a set of five questions was presented to participants to collect their response confidence, after which the next animation started. On average, participants took approximately 16 minutes to complete the main experiment with the animated stimuli. Before and after being shown the animated displays, each participant completed a Short Stress State Questionnaire (Helton 2004) (see Annex 1: *Short Stress State Questionnaire*). Finally, participants were asked to fill out a post-test questionnaire to collect background information and judgements about the difficulty of solving the task (see Annex 4: *Post-Test Questionnaire of Experiment I*) (Maggi et al. 2016). The results of

[13] Smartband by: http://bodymonitor.de/.
[14] Emotiv EPOC by: http://emotiv.com/.
[15] Tobii TX300 eye tracker by: http://www.tobii.com/.

participants' task performance, spatial abilities, and electrodermal and brain activities are provided in the next section.

5.4 Results

5.4.1 Response Accuracy

On average, participants detected the accelerating aircraft with a mean response accuracy of approximately 63% (SD=32.08). As shown in Figure 19, participants' response accuracy differed significantly across the two expertise groups and animation design types.

In general, novices performed the task less accurately than experts in both animation conditions. However, this difference was significant only for continuous animations ($F(1,17)=22.19$, $p<.000$). In this case, novices performed the task with only 27.78% (SD=29.55) correct responses. In addition, novice accuracy scores for the continuous animations differed significantly from the novice accuracy for semi-static animations ($F(1,17)=6.38$, $p<.022$). They performed the task better with semi-static animations (M=63.13%, SD=31.35) than with continuous animations (27.78%, SD=29.55).

Overall, experts responded with a mean accuracy close to 80% (M=81.44%, SD=15.51). They responded slightly better with semi-static animations (M=86%, SD=15.58) compared to continuous animations (M=78%, SD=15.08). However, this difference across animation design conditions was not significant.

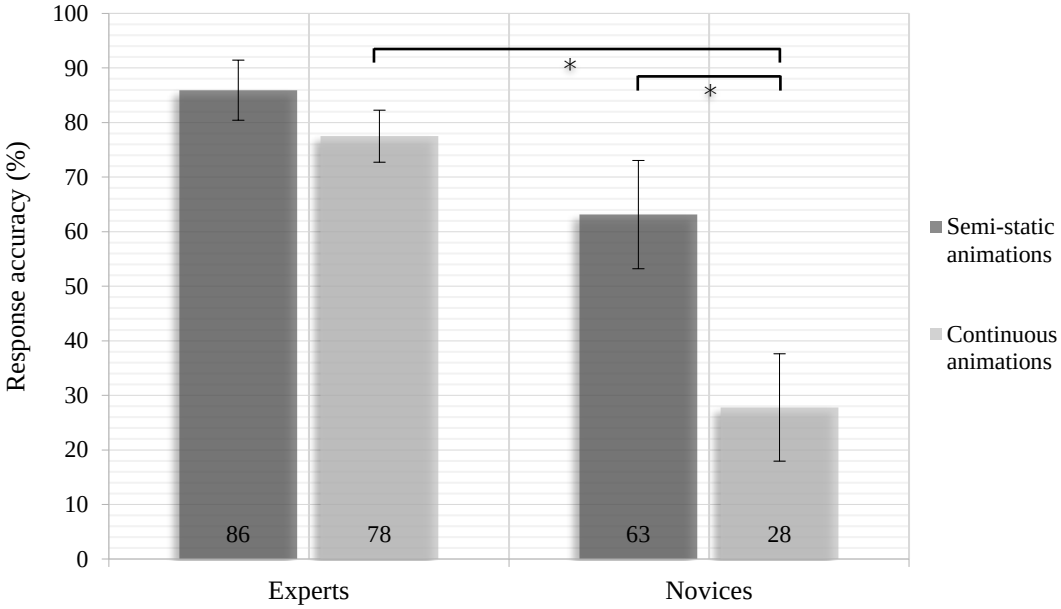


Figure 19: *Response accuracy for experts and novices across animation conditions (error bars show the standard error) (from Maggi & Fabrikant 2014a).*

Further, we analysed the effect of task complexity, i.e., first, the number of the depicted objects and second, differences of aircaft relative speeds, on participant task performance.

As shown in Figure 20, participant response accuracy for eight stimuli depicting four moving aircraft was compared with eight stimuli depicting eight moving aircraft. Overall, participants were more accurate in detecting the accelerating aircraft with displays depicting four moving aircraft (67.03%, SD=5.44), comapred to displays depicting eight aircraft (57.79%, SD= 5.86). This difference was significant ($F(37,1)$=4.90, p<.033).

Additionally, participant response accuracy differed significantly with respect to their expertise level, where experts performed significantly more accurately than novices in both task difficulty conditions (four objects: $F(37,1)$=17.69, p<.000; eight objects: $F(37,1)$=10.86, p<.002). Conversely, animation design affected participant response accuracy only in the test stimuli depicting four moving objects ($F(37,1)$=5.22, p<.028), and not with test stimuli depicting eight objects.

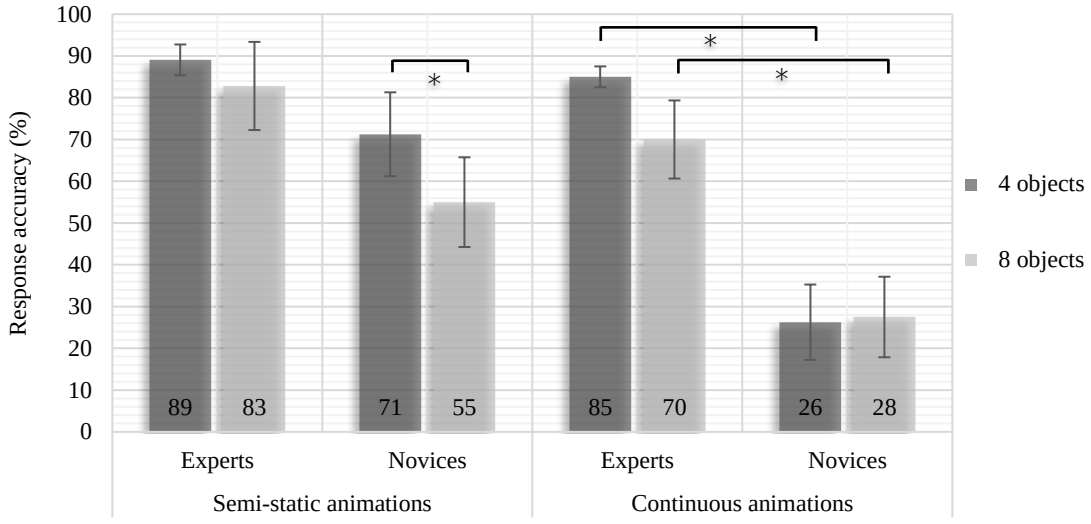


Figure 20: *Response accuracy for experts and novices across animation designs and number of depicted objects (error bars show standard error).*

Moreover, we found that differences in aircraft relative speeds might affect participant task performance, as well. We compared participant response accuracy for the four stimuli depicting aircraft moving at the same speed, except the target aircraft (ES condition) with their response accuracy for the twelve stimuli depicting aircraft moving at different speeds (DS condition). As expected, participants performed the task significantly more accurately with the same-speed stimuli (median (Mdn)=100%) compared with those depicting different-speed aircraft (Mdn=67%) (Z=-4.79, p<.000).

As shown in Figure 21, the response accuracy of the participants in both of the task difficulty conditions (i.e., ES and DS) is influenced by their expertise level as well. In general, experts performed the task more accurately than the novices in both of the task difficulty conditions (ES: Z=-2.65, p<.008; DS: Z=-2.97, p<.003). In addition, experts with continuous animations performed the task significantly more accurately in the ES condition than in the DS condition (Z=-2.69, p<.007). Novices with continuous animation conditions in the DS condition responded with an accuracy of 18.44% (SD=30.46, Mdn=8%) as this is barely above chance (i.e., 18.75%). This poor result significantly differed from their response accuracy with semi-static animations (Mdn=75%, Z=-2.27, p<.023). However, novice response accuracy did not differ significantly between the ES displays of both animation design types.

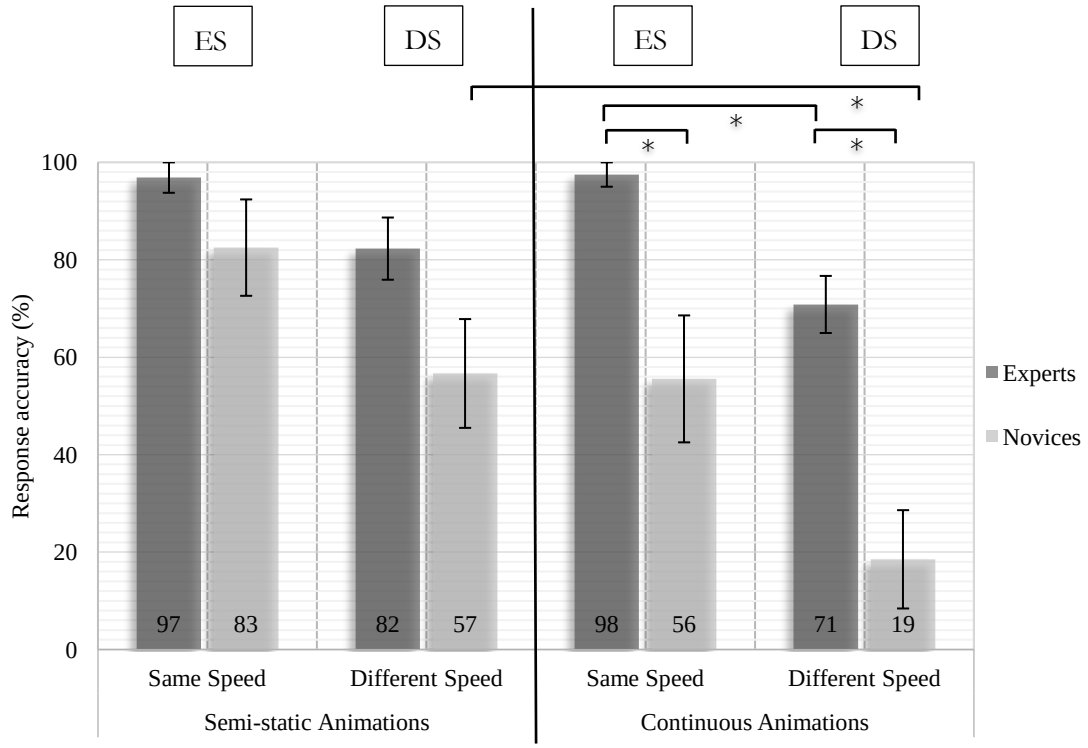


Figure 21: *Response accuracy for experts and novices in both animation design types and in both of the task difficulty conditions (same vs. different aircraft speeds; error bars show standard error).*

Finally, we further analysed participant responses in the continuous animation condition to better understand how they perceive perceptually more salient objects (i.e., faster moving aircraft) compared to thematically relevant objects (i.e., accelerating aircraft). In general, novices were more attracted by perceptual salient aircraft movements than experts with continuous animations. A total of 66.92% (SD=22.23) of novices and 21.67% (SD=15.86) of experts typically selected the perceptually more salient object, i.e., the fastest-moving object, rather than the task-relevant moving object, i.e., the accelerating aircraft. This result was significant ($F(1,22)$=32.95, p<.000).

5.4.2 **Response Time**

We also investigated whether animation design affects participant task efficiency (i.e., response time). We computed participant response time by considering only their task completion time for correct responses. On average, experts took less time to repond with semi-static animations (M=46.50 s, SD=9.73) than with continuous animations (M=61.00 s, SD=9.93). This difference is significant ($F(1,16)$=9.65, p<.007). In contrast, novice response time did not significantly change across animation design types, i.e., semi-static

animations (M=46.27 s, SD=16.69) compared to continuous animations (M=49.22 s, SD=20.51).

5.4.3 Eye Movements Analysis

The next step of our statistical analysis consisted in evaluating participants' eye movement data to better understand perceptual and cognitive processes that might explain their task performance. According to common eye movement metrics (Fabrikant et al. 2008), we systematically analysed which aircraft participants mostly focused on (i.e., eye fixation location), for how long (i.e., eye fixation duration), and how frequently (i.e., eye fixation rate) by means of an area of interest (AOI) analysis. Further, we assessed the influence of participant expertise level (i.e., experts vs novices) and animation design (i.e., semi-static animations vs continuous animations) on their strategies employed to solve the experiment task.

First, eye fixations have been filtered with respect to the 'I-VT' classification algorithm using a minimum fixation duration of 60 ms (Olsen 2012). Successively, all aircraft depicted on the test stimuli were assigned an AOIs to further measure eye fixation durations and eye fixation rates (i.e., the number of fixations per AOI, and per second). More specifically, we were interested in identifying participant search strategies, i.e. whether they were fixating perceptually on the more salient aircraft (i.e., the fastest aircraft), or the task-relevant aircraft (i.e., the accelerating aircraft). In the next section, I present the above-mentioned eye movement analysis. As response accuracy across expertise was only significant in continuous animations, the descriptive analysis focuses only on this animation design type. I excluded five of the 17 participants, as their eye movement data were of insufficient quality.

Descriptive Statistics

Overall, participants fixated on all aircraft for 484.82 ms (SD=119.90) in the continuous animation condition (Maggi et al. 2016). As expected, we found a significant fixation difference between experts and novices across task-relevant and perceptually more salient aircraft. As shown in Figure 22, experts fixated on both task-relevant (M=573.47 ms, SD=78.69) and perceptually more salient aircraft (M=497.15 ms, SD=57.82) for a significantly longer time than novices (task-relevant aircraft: M=400.95 ms, SD=128.71;

perceptually more salient aircraft: M=471.55 ms, SD=184.73). This result was significant ($F(1,11)$=6.60, p<.026). Moreover, experts fixated on task-relevant aircraft for a significantly longer time compared with the overall fixation duration ($t(12)$=4.07, p<.002). In contrast, novices exhibited significantly shorter fixation durations on task-relevant objects compared to the overall fixation duration ($t(12)$=-2.46, p<.032).

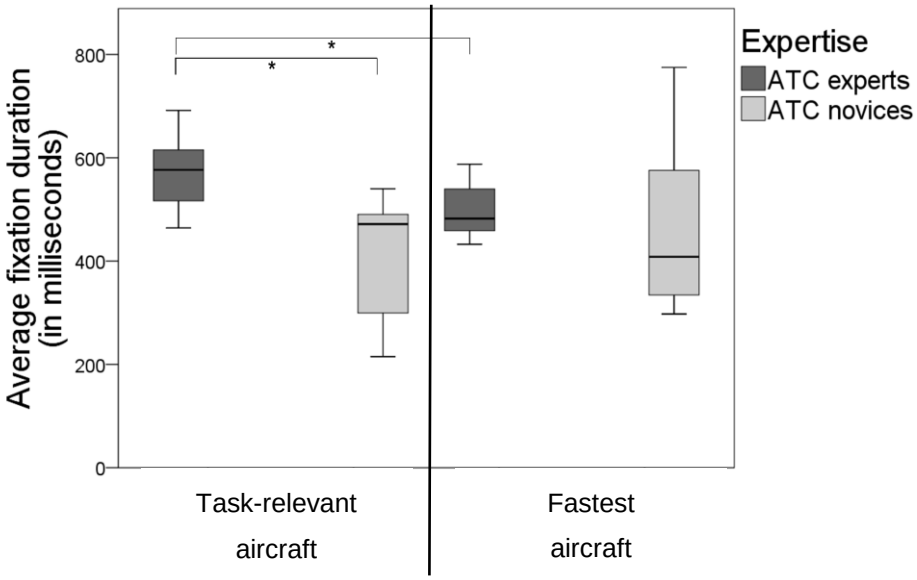


Figure 22: *Average fixation durations for AOIs in continuous animations, i.e., task-relevant (the two box plots on the left) and fastest (the two box plots on the right) aircraft, across expertise (from Maggi et al. 2016).*

The average eye fixation rate for all participants and all aircraft was 0.28 s^{-1} (SD=0.04). Overall, experts fixated on aircraft slightly more frequently (M=0.29 s^{-1}, SD=0.02) compared to novices (M=0.27 s^{-1}, SD=0.05). However, as shown in Figure 23, experts fixated on task-relevant objects significantly more frequently (M=0.37 s^{-1}, SD=0.028) than perceptually more salient ones (M=0.32 s^{-1}, SD=0.04) ($F(1,5)$=8.57, p<.033). Conversely, novices fixated on the fastest aircraft (M=0.38 s^{-1}, SD=0.08) significantly more often than the accelerating aircraft (M=0.26 s^{-1}, SD=0.11) ($F(1,4)$= 24.17, p<.008). Overall, experts fixated task-relevant aircraft more frequently than novices ($F(1,10)$=6.28, p<.031).

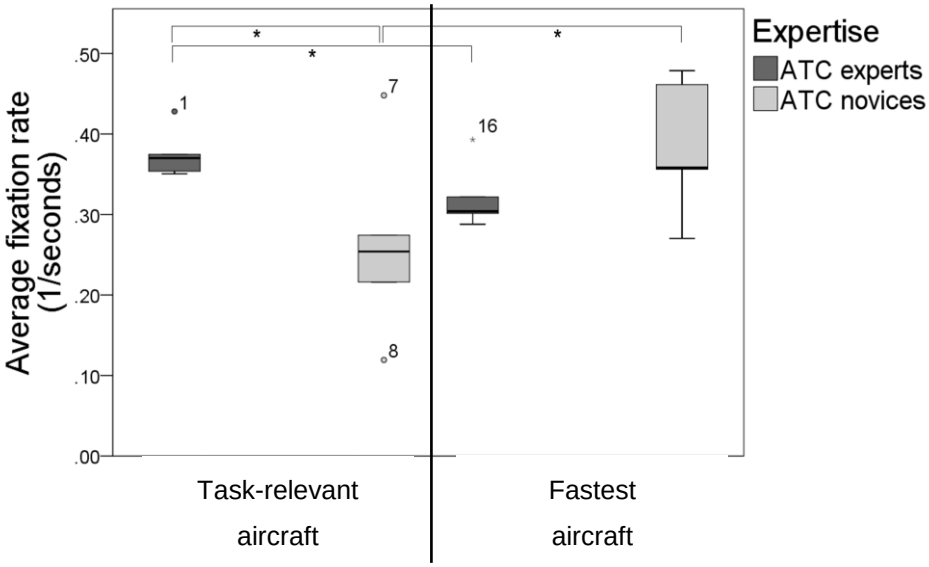


Figure 23: *Average fixation rate for AOIs in continuous animations, i.e., task-relevant (the two box plots on the left) and fastest (the two box plots on the right) aircraft, across expertise (from Maggi et al. 2016).*

Transition Matrices and Entropy Analysis

Along with the above-mentioned descriptive statistics, we further analysed eye movement data to investigate how participant eye fixation sequences between the depicted aircraft changed over time during an animation. This gaves us additional information about user visual search strategies. We thus calculated transition matrices of the recorded eye movement data across expertise, animation design, and task complexity (i.e., four versus eight aircraft), by using a R package (2016) developed by Krejtz et al. (2015), and adapted by Prof. Dr. Andrew Duchowski and myself. Transition matrices show the probability of the eye movement transitions to pass through a pair of AOIs on a visual scene, as well as the eye movement sequences within a specific AOI.

As shown in Figures 24 and 25, we first calculated the transition matrices of participant eye movement sequences of the animated displays depicting four aircraft. We assigned a specific AOI value to each of the four aircraft, i.e., AOI1, AOI2, AOI3, and AOI4. AOI1 always corresponded to the task-relevant aircraft (i.e., the accelerating aircraft), whereas AOI4 always corresponded to the perceptually more salient aircraft (i.e., the fastest aircraft), or the most peripheral and farthest object. We analysed eye movement transitions across expertise, animation design, and task difficulty (i.e., ES vs DS condition).

Figure 24 shows transition matrices of stimuli (S1 and S5) depicting four aircraft moving at the same speed (i.e., ES condition) across expertise and animation design. In all, we calculated four transition matrices for each condition: (A) experts with continuous animations, (B) experts with semi-static animations, (C) novices with continuous animations, and (D) novices with semi-static animations. In these transition matrices, AOI1 corresponds to the thematically relevant aircraft, and AOI4 to the most peripheral aircraft. Overall, we find that experts using semi-static animations are equally focused on both AOI1 and AOI4, but exhibit more frequent eye transitions compared to the other three conditions. Conversely, novice eye movement sequences with semi-static animations are more equally distributed on AOIs than experts with semi-static animations. Further, expert eye movement sequences are more frequent than those of novices with continuous animations. Novices with continuous animations made considerably fewer transitions between AOIs, and they were more focused on AOI1 and AOI4.

Figure 25 shows transition matrices of the stimuli (S1 and S5) depicting four aircraft moving at different speeds (i.e., DS condition) across expertise and animation design. Again, we calculated four transition matrices for each condition: (A) experts with continuous animations, (B) experts with semi-static animations, (C) novices with continuous animations, and (D) novices with semi-static animations. In these transition matrices, AOI1 corresponds to the acceleration aircraft, AOI4 corresponds to the fastest aircraft, whereas AOI3 is the most peripheral object. In general, we find that participants exhibit much less eye transitions between AOIs than in the ES condition. For example, experts' eye sequences with semi-static animations are more equally distributed on all AOIs, and much less frequent between AOIs compared to the same group in the ES condition. Conversely, experts with continuous animations are equally focused on both AOI1 and AOI4 (i.e., the fastest object), but exhibit more transitions compared with the other three groups. In addition, novices with continuous animations are more focused on the perceptual more salient objects (i.e., AOI4) compared to the other AOIs, and make few transitions between AOIs.

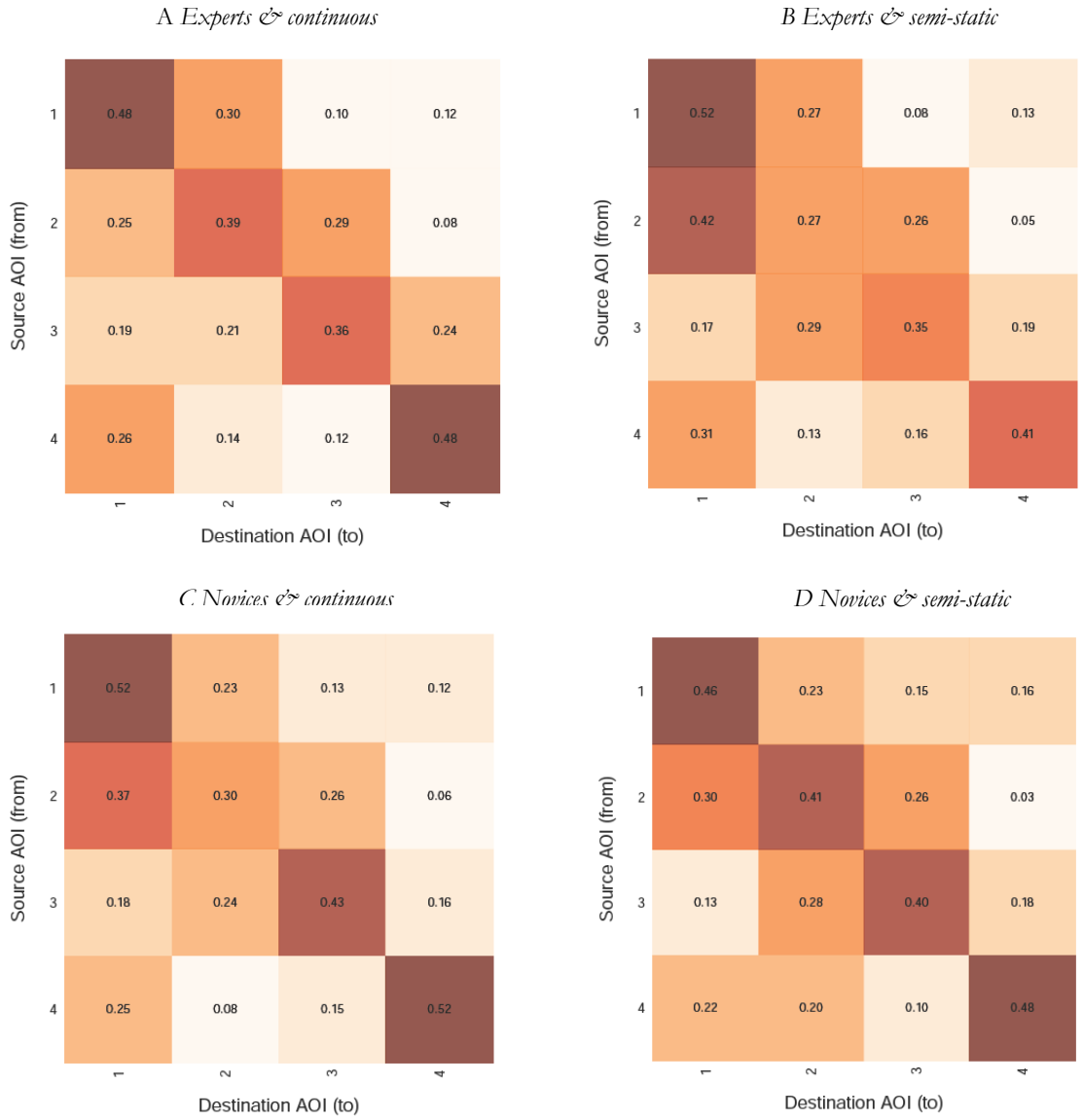


Figure 24: *Transition matrices of the stimuli (*S1 *and* S5*) depicting four objects moving at the same speed across expertise and animation design. Numbers from 1 to 4 correspond to the AOIs from* AOI1 *to* AOI4. AOI1 *is the thematically relevant object, and* AOI4 *is the most peripheral object. Each element of the matrix shows the probability that the eye saccades move from one AOI to another AOI on the screen, i.e., from the 'source AOI' to the 'destination AOI'. The different colour values highlight differences in the frequency of eye movement sequence (i.e., darker colours for more frequent transitions and lighter colours for less frequent transitions).*

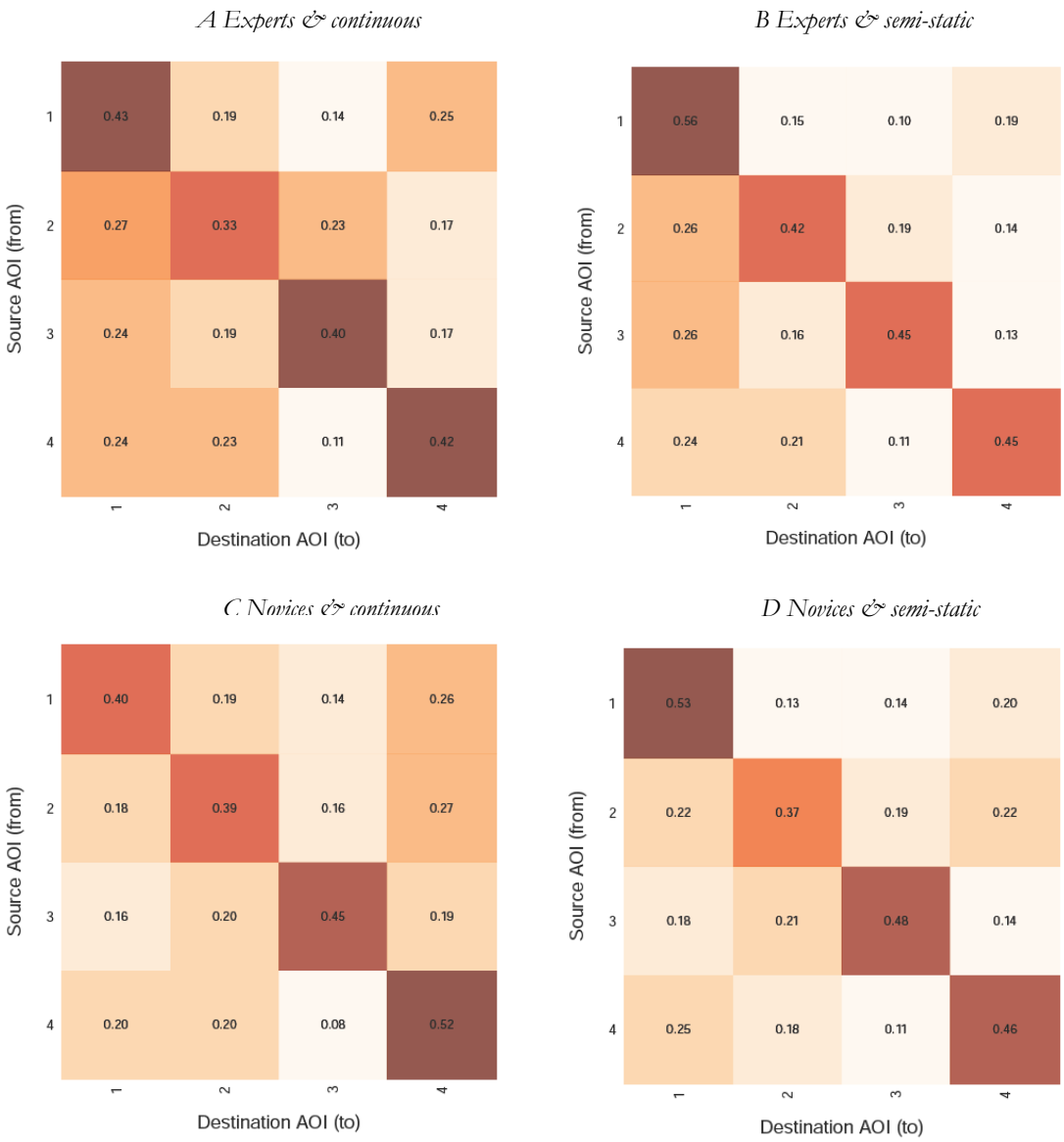


Figure 25: *Transition matrices of the stimuli (*S2–S4 *and* S6–S8*) depicting four objects moving at different speeds across expertise and animation design. Numbers from 1 to 4 correspond to the AOIs from* AOI1 *to* AOI4. AOI1 *is the thematically relevant aircraft, and* AOI4 *is the fastest aircraft. Each element of the matrix shows the probability that the eye saccades move from one AOI to another AOI on the screen, i.e., from the 'source AOI' to the 'destination AOI'. The different colour values highlight differences in the frequency of eye movement sequence (i.e., darker colours for more frequent transitions and lighter colours for less frequent transitions).*

The corresponding transition matrices of the animated displays depicting eight aircraft have been computed as well. AOI1 always corresponds to the task-relevant objects (i.e., the accelerating aircraft), whereas AOI7 and AOI8 always correspond to the more perceptually salient aircraft (i.e., the fastest aircraft). AOI3, AOI4 and AOI5 are the most peripheral and farthest object.

Figure 26 shows transition matrices of the stimuli (S9 and S13) depicting eight aircraft moving at equal speed (i.e., ES condition) across expertise and animation design. As for the animated displays depicting four aircraft, we calculated four transition matrices for each condition: (A) experts with continuous animations, (B) experts with semi-static animations, (C) novices with continuous animations, and (D) novices with semi-static animations.

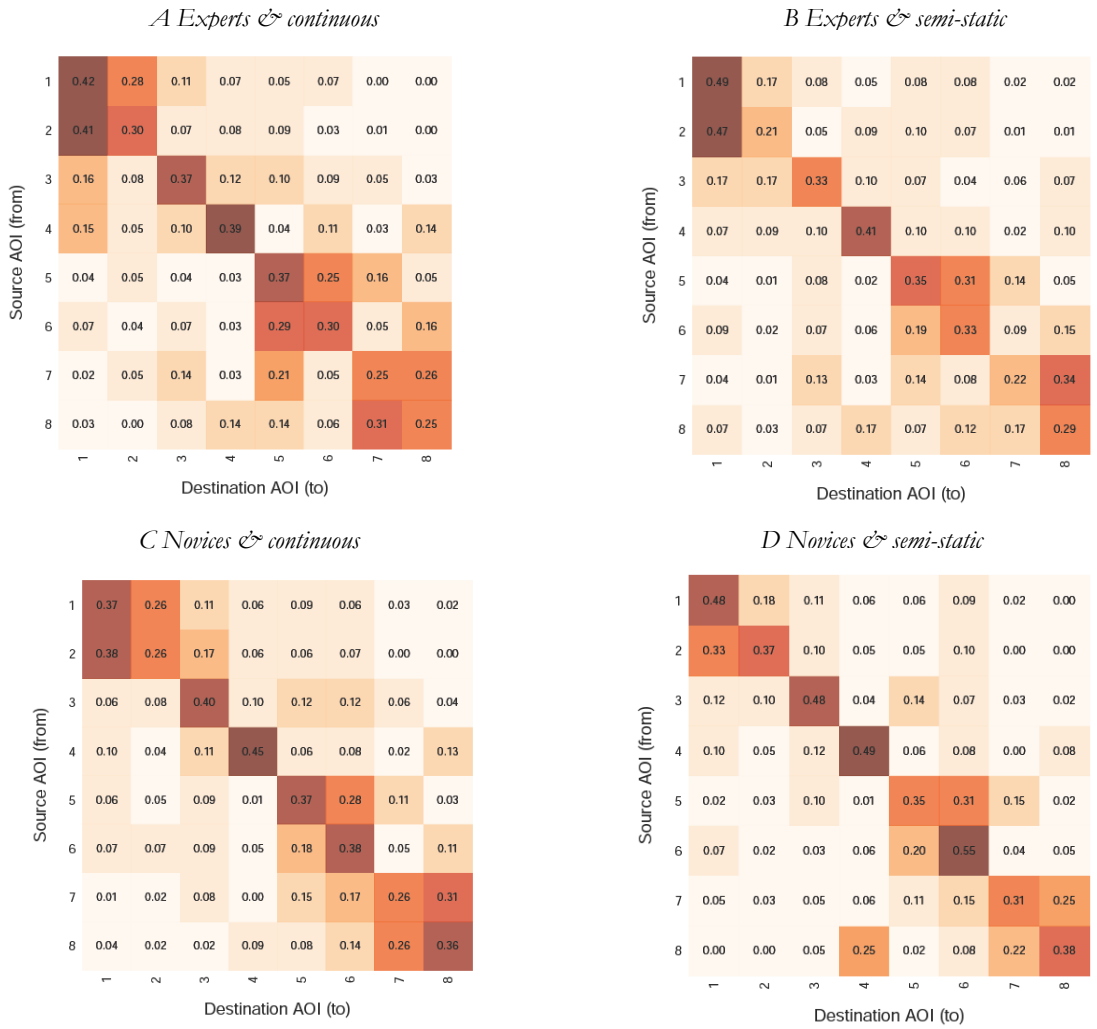


Figure 26: *Transition matrices of the stimuli (*S9 *and* S13*) depicting eight aircraft across expertise and animation design. Numbers from 1 to 8 correspond to the AOIs from* AOI1 *to* AOI8. AOI1 *is the thematically relevant aircraft;* AOI3, AOI4 *and* AOI5 *are the most peripheral aircraft. Each element of the matrix shows the probability that the eye saccades move from one AOI to another AOI on the screen, i.e., from the 'source AOI' to the 'destination AOI'. The different colour values highlight differences in the frequency of eye movement sequence (i.e., darker colours for more frequent transitions and lighter colours for less frequent transitions).*

In the transition matrices showed in Figure 26, AOI1 corresponds to the acceleration aircraft, whereas AOI3, AOI4 and AOI5 are the most peripheral aircraft. In addition,

aircraft that correspond to the following AOI-pair are more close to one another: AOI1 and AOI2, AOI5 and AOI6, and AOI7 and AOI8.

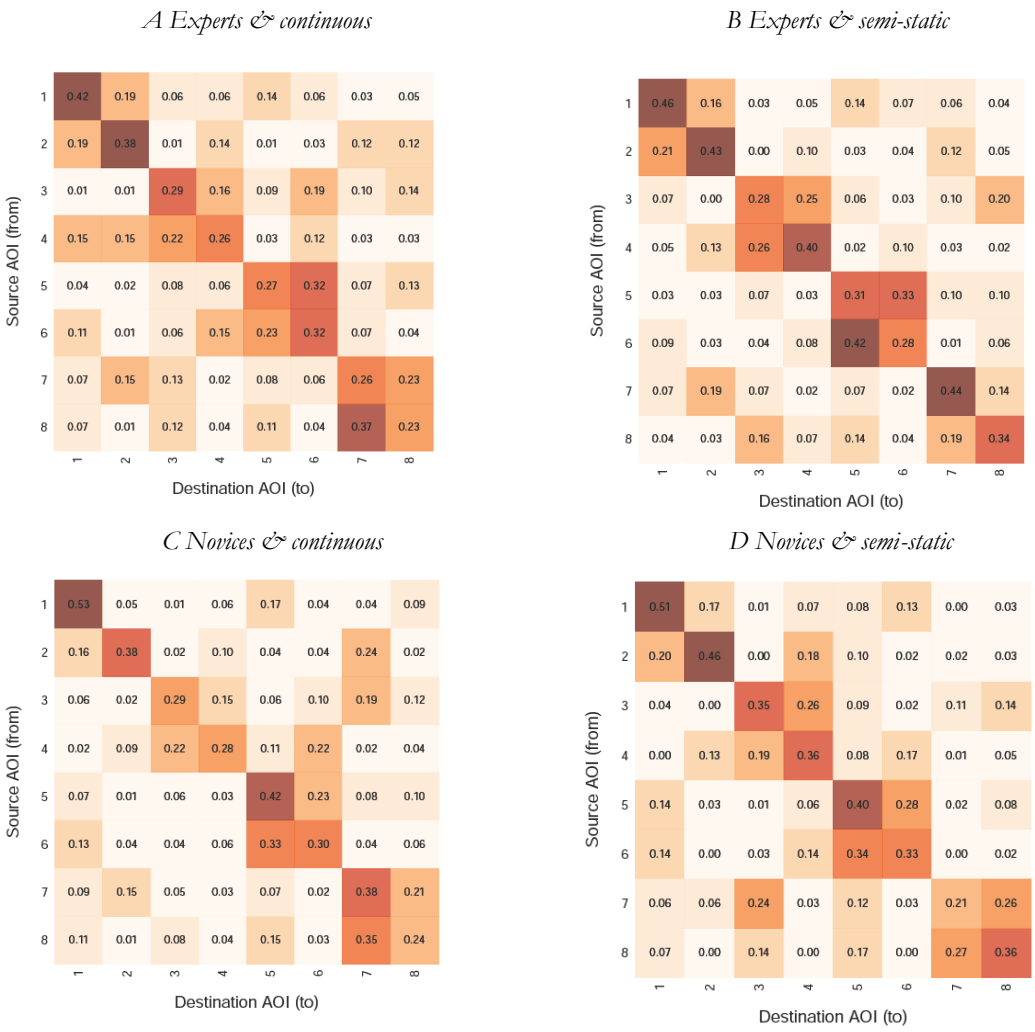


Figure 27: *Transition matrices of the stimuli (*S10–S12 *and* S14–S16*) depicting eight objects across expertise and animation design. Numbers from 1 to 8 correspond to the AOIs from* AOI1 *to* AOI8. AOI1 *is the thematically relevant aircraft;* AOI7 *and* AOI8 *correspond to the fastest aircraft;* AOI3, AOI4 *and* AOI5 *are the most peripheral aircraft. Each element of the matrix shows the probability that the eye saccades move from one AOI to another AOI on the screen, i.e., from the 'source AOI' to the 'destination AOI'. The different colour values highlight differences in the frequency of eye movement sequence (i.e., darker colours for more frequent transitions and lighter colours for less frequent transitions).*

In the ES condition (Figure 26), we find that overall participants made more transitions with continuous animations than with semi-static animations. In addition, transitions are more abundant between aircraft closer to each other (see clusters in Figure 26). Experts are more focused on the thematic relevant aircraft and on the most peripheral objects (i.e., AOI3, AOI4 and AOI5). Novices and experts using semi-static animations fixated the depicted aircraft in a similar, but novice eye movement sequences are less frequent than

those of experts. Conversely, the attention of novices using continuous animations is more focused on the most peripheral aircraft.

In the DS condition (Figure 27), AOI1 corresponds to the accelerating aircraft, AOI7 and AOI8 correspond to the fastest aircraft, whereas AOI3, AOI4 and AOI5 are the most peripheral aircraft. Overall, we find that participants focus their attention more on both thematic relevant information and on fastest and peripheral objects compared to other aircraft. Moreover, participant eye movement sequences occur more often between aircraft moving at similar speeds than other moving objects. Moreover, expert eye fixations on AOIs are more equally distributed than those of novices. Novices using continuous animations made fewer transitions than experts.

Transition and Stationary Entropy

We wish to further explore participant aircraft detection strategies by analysing their eye movement sequences. Standard eye sequence analyses are not suited for our data, due to their complexity. For this reason, we analysed eye movement data by means of *entropy metrics* (Krejtz et al. 2014). To analyse the collected eye movement data, we computed entropy using an R package (2016) developed by Krejtz et al. (2015) and adapted by Prof. Dr. Andrew Duchowski and myself. We calculated two different kinds of entropy metrics, i.e., *transition entropy (Ht)* and *stationary entropy (Hs)*. According to Krejtz et al. (2014), Ht corresponds to the complexity of eye movement sequences between AOIs, whereas Hs indicates the participants' interest focus on specific AOIs. High Ht values indicate frequent switching between AOIs. They suggest that participant eye movement sequences between AOIs are more complex and visual searches more exploratory. In contrast, low Ht values indicate a longer participant focus on specific AOIs, and frequent eye movement sequences between certain AOIs. High Hs values suggest also that participant visual attention between AOIs is more homogeneously distributed, whereas low Hs values indicate that participants concentrated their eye fixations more often on certain AOIs. With both entropy measures it is possible to make inferences about the complexity level of participant's eye sequences and where they concentrate mostly their eye fixations. This allow us, for example, to better understand whether participants use a target, or a more random, strategy for detecting the accelerating aircraft.

As shown in Figure 28 and Figure 29, we analysed the transition entropy (Ht) and the stationary entropy (Hs) of participant eye movement data across expertise and animation

design. Overall, we found that Ht and Hs differ significantly across the two expertise groups (Ht: U=31, p<.018; Hs: U=28.5, p<.011). However, we did not find any significant difference between the two animation design types.

We further calculated Ht and Hs of participant eye movement data for the animated displays depicting four aircraft, and across task difficulty levels (i.e., ES vs DS condition, and four vs. eight aircraft depicted). We found that the Ht values of participant eye movement data across ES and DS conditions differed significantly ($\chi^2(1)$=8.17, p<.004). The Hs values of the test stimuli depicting 4 objects moving at different speeds (S2, S3, S4, S6, S7, and S8) differed significantly between experts and novices ($\chi^2(1)$=4.56, p<.033).

Across the two task difficulty levels (i.e., four vs eight depicted aircraft), we found that the Ht values of the test stimuli depicting eight objects were much lower than the Ht values of the test stimuli depicting four objects. This means that the eye movement sequences of the participants were more predictable by eight aircraft than by four aircraft. Moreover, experts with test stimuli depicting 8 objects were much less focused on AOIs than with test stimuli depicting 4 objects. This might indicate that experts employed different strategies according to the task complexity level. With four objects, experts completed more transitions between AOIs and they were less focused on a certain AOI. Conversely, with eight objects, experts were more inclined to compare objects with similar speeds. However, in contrast novices fixated not only on objects closer to one another, but also on individual aircraft moving farther than other aircraft. In addition, novices looking at continuous animations engaged in fewer transitions, and were more focused on certain AOIs (fastest or peripheral objects) than novices with semi-static animations and experts using both animation designs.

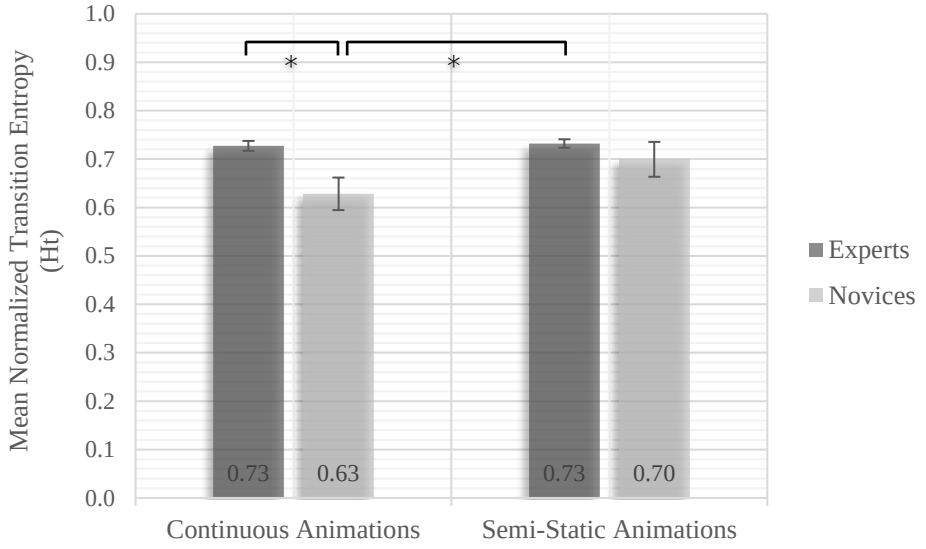


Figure 28: *Mean normalized transition entropy (Ht) of all stimuli across animation design and expertise.*

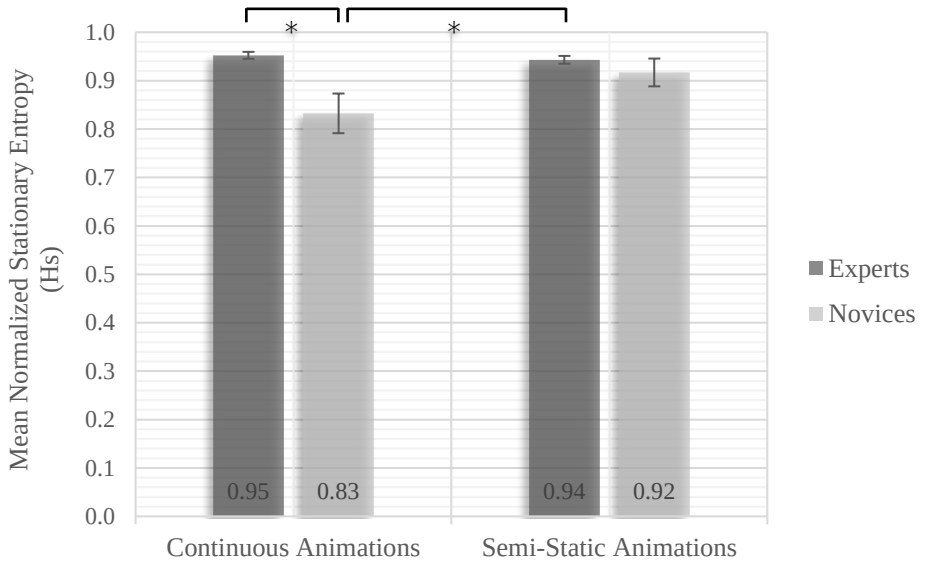


Figure 29: *Mean normalized stationary entropy (Hs) of all stimuli across animation design and expertise.*

We now present our results concerning the influence of users-related factors on the effectiveness of the detection task.

5.4.4 User-Related Factors

Self-Confidence

We collected participants' self-confidence scores by means of a questionnaire administered after each of the 16 tested stimuli. We accordingly collected 16 self-confidence judgements per participant, and we subsequently analysed their mean scores across animation design type, expertise level, and gender. On average, participants judged themselves with an average score of 3.78 (SE = 0.05) on a Likert scale of 1–5 (1= not confident at all, 5 = very confident). Their confidence scores did not vary significantly across animation design types, expertise level, or gender. We further analysed participant spatial abilities, collected by means of the Hidden Pattern Test, and their influence on task performance. The corresponding results are presented in the next section.

Spatial Abilities

We analysed participant' spatial abilities across expertise and animation design. Overall, results of the Hidden Pattern Test (HPT) showed a significant difference in participant spatial abilities across expertise ($F(1,35)$=14.61, p<.001). Experts solved the HPT test with 66.94% (SD=11.48) correct responses, while novices had a response accuracy of 48.05% (SD=17.74).

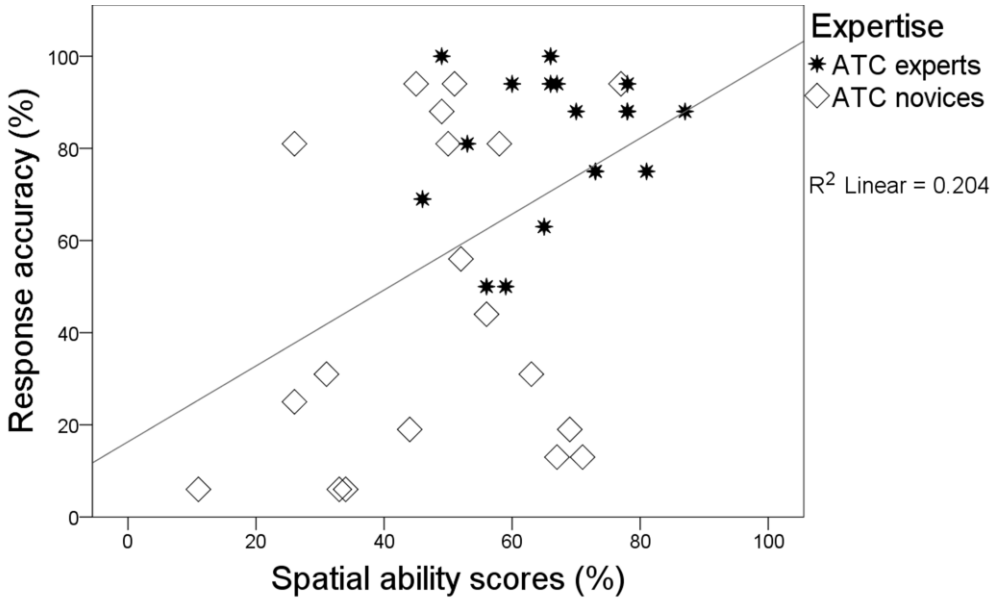


Figure 30: *Correlation of participants' response accuracy and spatial abilities scores across expertise (stars indicate experts, and diamonds signify novices) (from Maggi et al. 2016).*

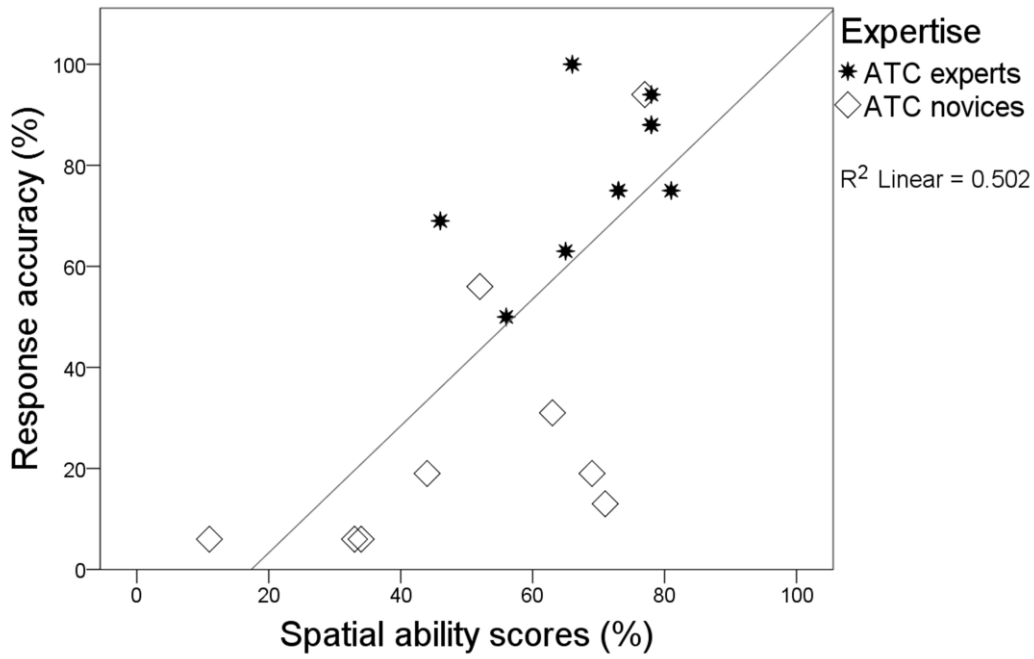


Figure 31: *Correlation of participants' response accuracy and spatial ability scores for the continuous displays (stars indicate experts, and diamonds signify novices) (from Maggi et al. 2016).*

Spatial abilities of the participants were also compared across the two animation types, but we did not find any significant differences. The Pearson correlation coefficient between participant response accuracy and their spatial ability scores for the semi-static animation condition indicates a moderate positive correlation between these two variables (r=0.452, N=37, p<.005), as shown in Figure 30. In the continuous animation condition (Figure 31), this relationship is stronger (r=0.71, N=19, p<.001). Besides their spatial abilities, we further analysed participant affective states during the experiment by analysing their responses to the Short Stress Questionnaire (SSSQ; Helton 2004).

Stress Related-Factors

To make inferences about participants' affective states during the experiment, we analysed the subjective stress-related factors (i.e., engagement, distress, and worry) captured with the SSSQ across animation design and expertise. On average, the overall results showed that experts were less engaged than novices (Figure 32). However, this result was not significant.

We further analysed the SSSQ change z-scores of experts and novices across animation designs. As shown in Figure 33a, the SSSQ change z-scores of experts showed more engagement, less distress, and less worry with the familiar semi-static animations,

compared with novices (Maggi & Fabrikant 2014a). However, only the engagement scores were significantly different between experts and novices ($t(16)=3.00$, $p<.050$). Distress and worry were not significantly different across the two expertise groups, probably because of the large variance in the novice population.

Conversely, as shown in Figure 33b, with the continuous animations experts and novices showed similar low engagement patterns, and novices exhibited more distress and worry than experts. However, none of the three SSSQ scores showed significant differences. This might again be due to the large variance in the novice group (e.g., distress variable).

Besides self-reported stress-related factors, we further analysed participant skin conductance arousals as well as brain activity, which might give us more insight to their affective states, motivation and cognitive load when solving the detection task. The analysis of the affective states and cognitive processes of participants in user studies by means of EDA and EEG data is a novel approach in cartography and GIScience research.

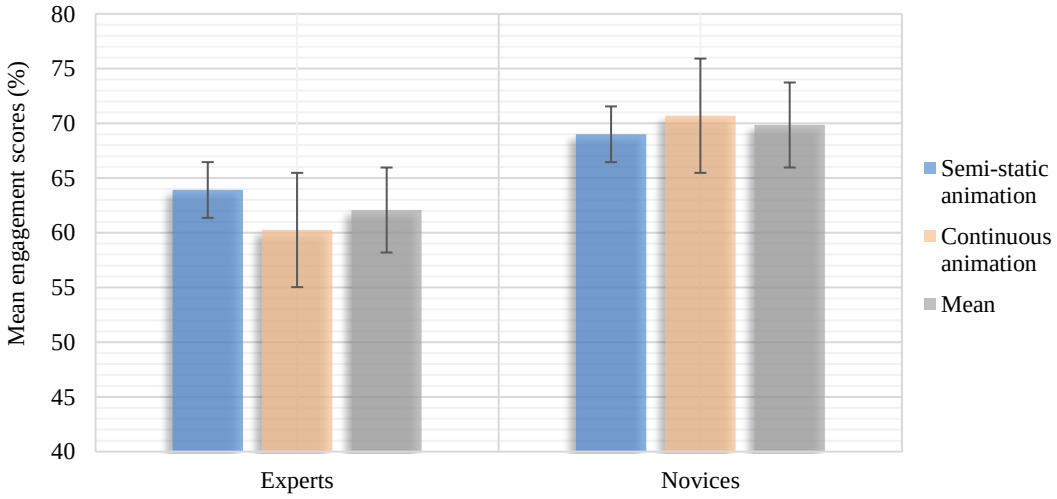
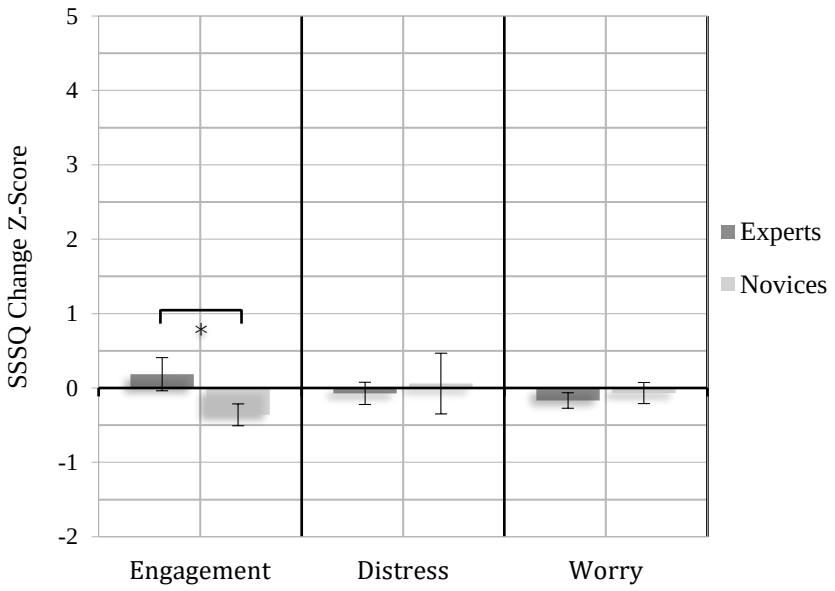
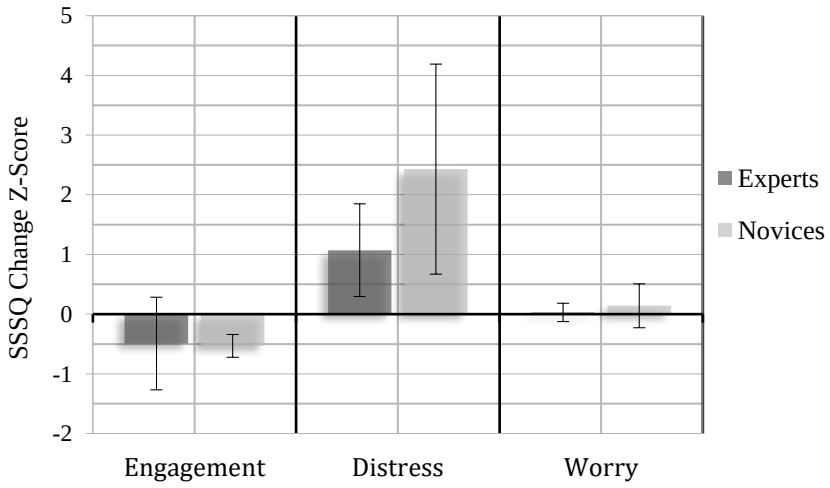


Figure 32: *Mean engagement scores across expertise and animation design.*



(a) Semi-static animation.



(b) Continuous animation.

Figure 33: *SSSQ scores of participants for the (a) semi-static animations and participants in the (b) continuous animations (from Maggi & Fabrikant 2014).*

Electrodermal Activity (EDA)

Besides self-reported SSSQ scores, we also analysed participants' skin conductance responses, recorded with the Smartband sensor device, across animation design and expertise according to standard metrics (Boucsein 1992; Lykken 1972). Data have been computed using the linear mixed-effects model. This approach is better than GLM to analyse complex repeated measurements of participants. According to a fixed-effects model, we calculated the effect of animation design and expertise on participants' emotional state. However, we were able to only analyse a subset of 20 participants (13 experts and 7 novices), due to noise in the recorded EDA data. Overall, experts exhibit higher arousal levels (i.e., based on mean area bounded by the SCR curve) than novices, as shown in Figure 34. However, this EDA difference between experts and novices is only significant for the semi-static animations, due to large variances in the continuous animation condition.

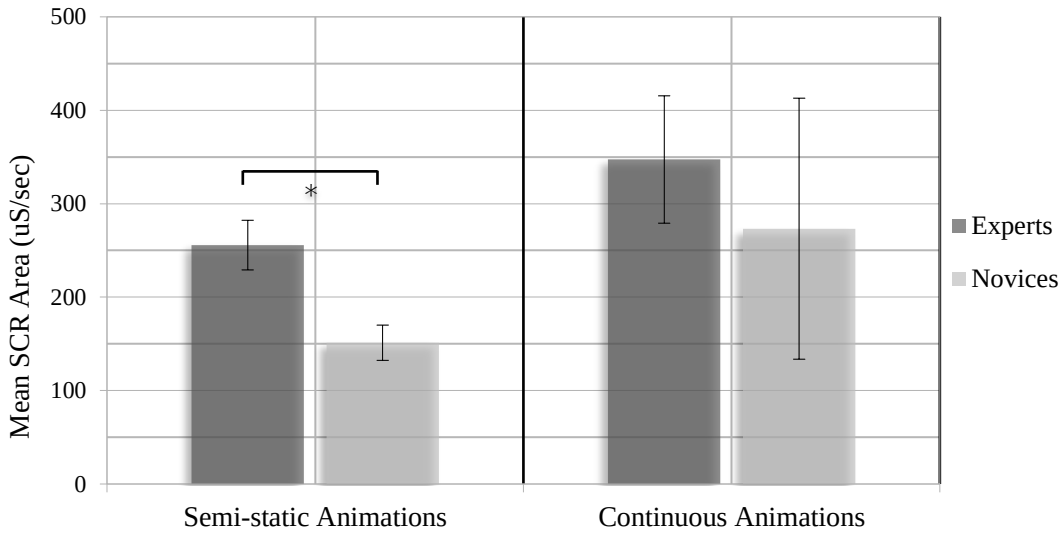


Figure 34: *Participants' electrodermal intensity, with respect to the mean area bounded by the SCR curve, across expertise and animation design (Maggi, Fabrikant 2014a).*

Analysis of Cognitive Load and Motivation from the EEG Data

We additionally analysed participant brain activity to determine motivation level and cognitive load while performing the experiment to better explain their task performance. To measure motivation, as explained in Section 4.1.3, *Measuring Brain Activity*, we performed a *Frontal Alpha Asymmetry* (FAA) analysis with recorded EEG data. First, the collected EEG data were cleaned and filtered following a standard procedure. We then extracted alpha power values using the *EEGLAB*[16] tool box, and then computed FAA scores for each participant using the *STLab 0.2.4*[17] tool box, both available in MATLAB (Briesemeister et al. 2013).

Overall, the averaged *FAA score* across participants was 0.62 (SD=2.88). Novices showed significantly higher FAA scores than experts ($F(1,23)$=5.86, p<.024) (Figure 35). According to Briesemeister et al. (2013), positive FFA scores relate to larger relative right-hemispheric activation, which in turn is an indicator of approach-related motivation and positive affective states, such as curiosity and confidence. Interestingly, experts exhibited larger right-hemispheric activation suggesting withdrawal-related motivation, which in turn indicates more negative affective states (e.g., boredom, lack of interest, fear, or overload). We did not find any significant difference in the FAA scores between semi-static and continuous animations, nor between test stimuli displaying four aircraft and test stimuli displaying eight aircraft. However, with continuous animations, participants showed higher FFA values than with semi-static animations.

We then assessed overall *cognitive load* of participants during the experiment by calculating mean alpha power values from the collected EEG data. Similarly to the FAA analysis, we found a significant difference in the alpha power values between experts and novices (F(1,23)=5.73, p<.026), but not across animation conditions, as shown in Figure 36. Novices showed a significant lower alpha power than experts. According to Antonenko et al. (2010), alpha power tends to increase as cognitive load decreases. Furthermore, an increase of alpha power may also reflect visual information processing guided by top-down mechanisms rather than of bottom-up processing demands (Benedek et al. 2014). This means that the cognitive load of our tested novices was higher than that of the experts and that novices were guided by bottom-up processes rather than by top-down processes.

---

[16] EEGLAB: http://sccn.ucsd.edu/eeglab/.

[17] STLab 0.2.4 has been developed by Briesemeister et al. 2013 to calculate FAA scores.

Further, novices in the continuous animation condition exhibited a slight negative trend, whereas experts had an overall higher positive value. Negative alpha power values indicate a decrease of the alpha power and thus an increase of the mental effort. This probably means that novices had more difficulties in solving the task than experts and novices in the semi-static animation condition.

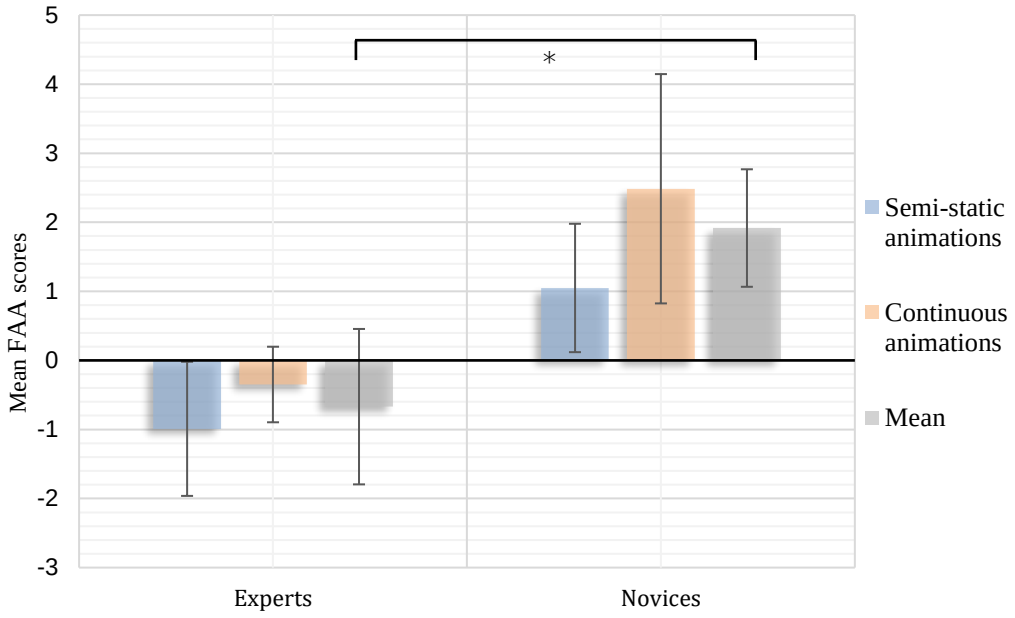


Figure 35: *FAA scores of experts and novices across the two animation design conditions, i.e., semi-static and continuous animations.*

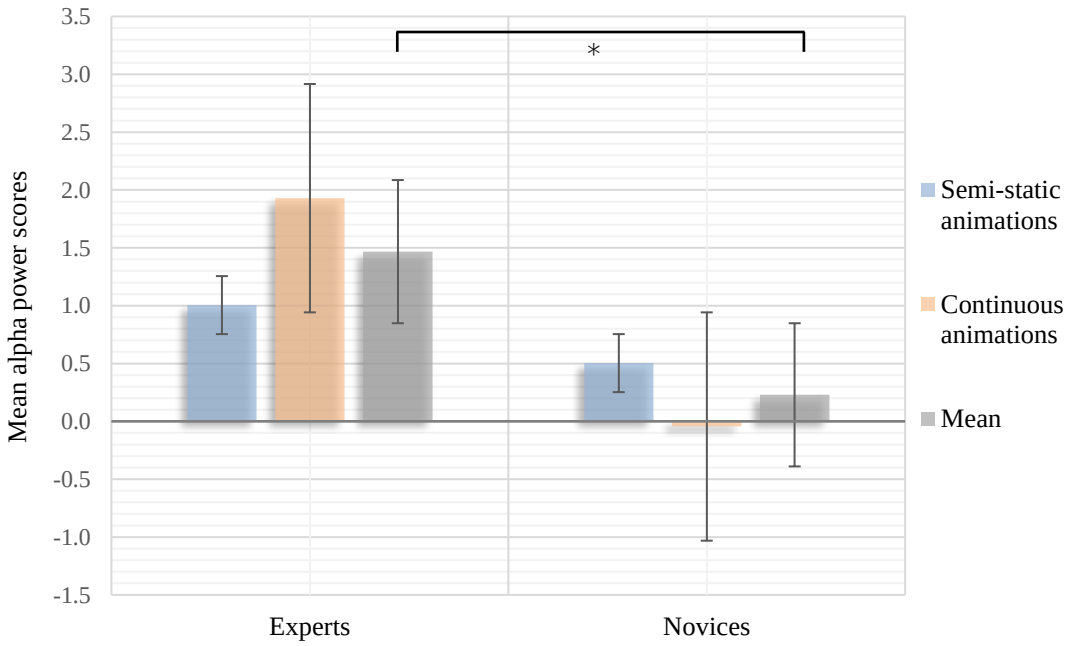


Figure 36: *Mean alpha power scores across animation design types and expertise.*

5.5 Key Findings and Discussion of Experiment I

In this first user study, we investigated how *task difficulty* (i.e., number of the depicted objects, and differences in their relative speeds), *animation design* (i.e., rate of change or smoothness of the transitions between display scenes), and *user-related factors* (i.e., perceptual, cognitive, and affective processes) might influence the apprehension of movement changes (i.e., accelerations) with ATC animated displays (cf. Section 5.2, *Specific Research Questions and Working Hypotheses*).

Our results suggest that the use context and task (animations depicting 4 vs 8 aircraft, and with aircraft moving at similar vs different relative speeds), animation design (semi-static vs continuous animations), and user characteristics (spatial abilities, expertise, and stress factors) do indeed influence visuospatial inference and decision-making with animated displays. In the next sections, I detail key findings of these three factors separately.

5.5.1 Context of Use and Task

We tested the effect of two task complexity factors on the apprehension of movement changes (i.e., aircraft accelerations) with animations: the number of the depicted objects (animations depicting 4 vs. 8 aircraft) and differences in their relative motion (same vs. different relative speeds). The task was designed according to the first SA principle (Endsley 1995; Grier 2015), i.e., how effectively viewers apprehend aircraft movement changes, and with realistic ATC displays. Our hypothesis, mentioned in Section 5.2, *Specific Research Questions and Working Hypotheses,* is as follows:

> ***WH 1:*** *Users can effectively apprehend movement changes if few elements, that also move at similar relative speed, are visualized simultaneously (Koffka 1935; Ware 2013). Increasing the number of moving objects displayed on the screen, as well as the difference of their relative speed, negatively influences the apprehension of movement changes.*

Our results show that the *number of the depicted objects* indeed significantly influenced the apprehension of movement changes (i.e., detection of aircraft accelerations) in animated displays, as we hypothesized (cf. **WH 1**). In general, participants (both experts and novices) tended to detect movement changes less accurately with displays depicting 8 objects compared to displays depicting only 4 objects. This result is in line with previous animation

studies. Pylyshyn and Storm (1988), as well as Ware (2013), found that users are able to keep track of a maximum of 4 or 5 objects simultaneously for several seconds.

However, the number of moving objects depicted significantly influenced the correct apprehension of movement changes depending on the animation design and on user expertise. More specifically, we found that this effect is significant only for novices with semi-static animations. Experts, in the semi-static and continuous animation conditions, showed similar patterns, and their performance across displays depicting 4 objects or 8 objects did not significantly differ. It seems that the findings presented by Ware (2013) are especially valid for novices, but not for experts. This difference between experts and novices may be due to training and to different visual searching strategies developed to effectively detect movement changes. On average, experts are trained to monitor 8–12 aircraft simultaneously, and even, in large airports, over 20 aircraft at the same time. Niessen et al. (1999) explain that air traffic controllers are able to track several objects simultaneously by giving priority to the most relevant tasks. Moreover, Cavanagh and Alvarez (2005) claim that users can track more than 4 objects at the same time by grouping multiple objects that share some common motion patterns. This mental process of grouping similar movement patterns might be explained by the Gestalt theory of common fate (Koffka 1935). It states that similar movement patterns (e.g., objects moving at the same direction or speed) are perceived as being more related compared to different movement patterns. For this reason, they are grouped together and mentally processed as a single entity. This might reduce the cognitive effort of processing more than 4 objects simultaneously.

From the entropy analysis that we performed on the eye movement behaviour, we found that both transition and stationary entropies of experts are significantly higher than those of novices. In addition, both transition and stationary entropies with displays depicting 4 objects were significantly higher than displays depicting 8 objects. According to Krejtz et al. (2014), this means that the eye movement sequence patterns (i.e., eye movement switching between AOIs) of novices, and of all participants looking at displays depicting 8 objects, seem to be less complex. In addition, their eye fixations are more focused on the same areas than are those of the experts and in the displays with 4 objects. In other words, it seems that novices focus more frequently on moving objects with similar motion patterns and objects that are closer to one another. Experts pay attention not only to all objects depicted in the scene, but also to those objects that are not close, or objects

moving at different speeds. This finding is confirmed by prior ATC research (e.g., Wang et al. 2015; McClung & Kang 2016), which highlighted that air traffic controllers' visual scanning strategies aim at finding conflicts or anomalous aircraft movement patterns, and are more effective than those of novices; for example, air traffic controllers repeatedly scan the whole screen using circular eye movement patterns. In addition, Stein (1989) found that the saccades velocity of experts is higher than those of novices, meaning that their eye movement sequences are more complex and less focused on specific objects depicted on the scene.

Furthermore, novices in the continuous animation condition performed similarly with both displays depicting 4 and 8 aircraft, but the accuracy of their answers was only about 27%. From the eye movement and entropy analysis, we found that they are more focused on the more perceptually salient moving objects (i.e., the fastest aircraft) than on the thematically relevant ones (i.e., accelerating aircraft), and this was true irrespectively of the number of the depicted objects. This is consistent with our hypothesis that experts are more guided by top-down processes, while novices are guided more by bottom-up processes in visuospatial detection tasks, as explained in Section 5.4.4, *User-Related Factors*.

Differences in the relative speed of the depicted objects (i.e., aircraft moving at the same or at different speeds) also influenced user information processing. As expected (cf. **WH 1**), users performed better with the displays depicting aircraft moving at the same speed than with displays depicting aircraft moving at different speeds. This may have happened because of users' pre-attentive processing of visuospatial features (Ware 2013). Anomalous movement patterns embedded in a homogeneous context are pre-attentively distinguishable because they pop from their surroundings. Similarly, an accelerating aircraft surrounded by aircraft moving all at the same speeds should be more easily detectable than with aircraft moving at different speeds.

Differences of relative speeds between moving objects affected participants in a similar way, irrespectively of their expertise level. Both experts and novices performed more accurately with displays showing aircraft with the same speed than with displays showing aircraft at different speeds. However, the response accuracies of novices were additionally influenced by the animation design type. Novices with semi-static animations performed more accurately than novices with continuous animations in displays showing aircraft at different speeds. In contrast, experts performed more accurately with displays showing

aircraft with the same speed than with displays showing aircraft at different speeds, independently of the animation design type. The influence of the animation design and user characteristics on the apprehension of movement changes with animated displays is discussed in more detail in the next two sections.

5.5.2 Animation Design

We tested the effect of animation design on the apprehension of movement changes (i.e., aircraft accelerations) by manipulating the rate of change of the animated displays, i.e., by ¼ Hz (one display every four seconds) vs 60 Hz (60 displays every second). Our hypothesis concerning animation design, mentioned in Section 5.2, *Specific Research Questions and Working Hypotheses*, is as follows:

> ***WH 2:*** *Users can more easily detect movement changes with continuous animations than with semi-static animations (Tversky et al. 2002).*

According to our above-mentioned hypothesis (cf. **WH 2**), we expected that users might more efficiently apprehend spatio-temporal data with continuous animations than with semi-static ones, because movement changes are congruently depicted to human mental models of movement changes over time, conforming to the *Congruence Principle of Good Representation* (Tversky et al. 2002). We found that animation design indeed influences the perception of movement changes in a visuospatial detection task with air traffic control displays, but, contrary to our expectations, participants performed the acceleration detection task more effectively with semi-static animations than with continuous animations.

In the past geovisualization literature (Garlandini & Fabrikant 2009; Goldsberry & Battersby 2009), transitions between dynamic scenes (i.e., *tweening*) are often recommended to mitigate the *change blindness effect*, such as in animated thematic maps. However, as Fabrikant (2005) pointed out, if the transition is too smooth, users might find difficulties in identifying slight, but relevant changes. In our experiment, the worst performance with continuous animations may thus be due to the *smooth and continuous transitions between scenes*, which make the perception of micro-step changes in aircraft speed more difficult compared with abruptly refreshing scenes (Fabrikant et al. 2010). It might be that speed

changes, refreshing continuously, were too smooth, and thus required a higher perceptual and cognitive demand on the working memory of users to be effectively detected.

In contrast, speed changes of semi-static animations are not continuously perceivable, but are seen as a sequence of discrete steps over time. Viewers have to infer speed differences and accelerations through the graphic encoding by means of different aircraft shapes; participants can identify the fastest-moving aircraft because of their elongated form. Since speed changes are abruptly refreshed every 4 seconds, the magnitude of change per time unit is greater than in continuous animations. These discrete and visually more salient changes might have facilitated the participants in their detection task, as demonstrated by previous change detection studies in psychology (Goddard & Clifford 2013) and predicted by the *Event Perception Theory* (Shipley & Zacks 2008).

However, for this experiment, we only tested speed changes, which in the ATC context occur very slowly. After the experiment, some experts reported that continuous animations might be more effective to detect *direction changes*, because they are more noticeable compared to aircraft speed changes. The continuous transitions between aircraft locations might be thus more useful for them when they have to continuously track aircraft direction changes, or to better apprehend the relative motion between multiple aircraft. This is what we investigated further in Chapter 6, *Experiment II*.

In addition, the animation design type influences predominantly novices. Experts performed well with both animation designs, without significant differences in their response accuracies between semi-static and continuous animations. The only difference in the detection of movement changes between these two animation conditions is in their response efficiency (i.e., response time). Experts took slightly more time to detect the thematically relevant information with continuous animations than with the more familiar semi-static animations. That experts are slightly worse in continuous case might suggest they are not using the motion information to apprehend movement changes, but only the visual information of the ODS comet form. Novices may be trying to use motion information, but with less success. Consequently, our hypothesis (cf. **WH 2**) holds true only partly for novices. This difference between experts and novices, as well as other potential user-related factors influencing tasks performance, are discussed in more detail in the next section.

5.5.3 User Behaviour

As discussed in the previous section, we could see that novices solved the detection task less well with continuous animations than with semi-static animations. In contrast, experts performed well with both animation design types. Let us inspect expert and novice behaviour more closely.

Differences in task performance between novices and experts might be due to various factors. Our three hypotheses concerning user-factors, and mentioned in Section 5.2, *Specific Research Questions and Working Hypotheses*, are as follows:

> ***WH 3.1:*** *Experts will perform tasks more effectively and efficiently with both animation types than novices due to their previous training (Hegarty et al. 2003).*
>
> ***WH 3.2:*** *Experts might perform worse on tests with continuous animations because of their training being limited to semi-static animations (mismatch with controllers' heuristics, with reference to Lee & Klippel 2005). In general, users will perform tasks more effectively and efficiently with semi-static animations if they have high spatial abilities. Respectively, users will perform tasks more effectively and efficiently with the continuous design if they have low spatial abilities (Hegarty et al. 2003).*
>
> ***WH 3.3:*** *The type of animation with which users are asked to solve the required task, user expertise level and familiarity with the task might influence their emotional and affective states, and consequently their task performance.*

Before conducting our experiment, we hypothesised (cf. **WH 3.1**) that experts perform the tasks more effectively and efficiently than novices with both animation types due to their *previous knowledge, familiarity, and training* with the tested displays. In addition, our fourth hypothesis (cf. **WH 3.2**) is that experts would perform less well with continuous animations than with the familiar semi-static animations because of their training being limited to the semi-static design (Lee & Klippel 2005). Contrarily to our expectations, we found that experts performed more accurately than novices (at least 80% of correct responses) with both animation design types. Familiarity due to their training with semi-static animations allows them to effectively detect the accelerating aircraft despite the perceptual salience of thematically irrelevant information (i.e., fastest moving aircraft) with continuous animations. This finding highlights the importance of top-down mechanisms

for experts in information processing performance with both novel and familiar displays. In contrast, for novices bottom-up mechanisms seem to be more relevant than for experts. Our results concerning participant brain activity confirms that as well. The higher alpha power values of experts might indicate the activation of cognitive mechanisms related to more internally oriented attention (i.e., top-down activity) than bottom-up processes (Benedek et al. 2014; Fink & Benedek 2014).

The above-mentioned findings might be explained with previous studies on animations (Boucheix & Lowe 2010; Fabrikant 2005). These studies emphasized the importance of perceptual salient information in the decisional processes of novices. Users, especially novices, might process information according to perceptually salient features rather than to thematically relevant ones. According to the *Gestalt principle* of common fate (Koffka 1935), salient perceptual features pop out from the graphical display because of their dynamic contrast strengths with the information background. They are thus more easily distinguishable from their surroundings (Boucheix & Lowe 2010). Another explanation is that novices with continuous animations might concentrate their attention on perceptually salient objects because task-relevant information might be too difficult to be correctly perceived, as it requires too much perceptual and cognitive processing demand. Users might unconsciously and naïvely use this strategy to decrease their mental overload (Boucheix & Lowe 2010).

This outcome identified a correspondence with the eye movement analysis of participants with the continuous animations. Contrarily to experts, novices fixated more frequently upon the fastest aircraft than upon the accelerating aircraft. According to previous studies on eye movements (Holmqvist 2011), a higher eye fixation rate on specific scene objects indicates a higher noticeability of that area. In addition, eye fixation durations of novices are generally shorter than those of experts. Shorter fixations in continuous animations lead to reduced performance and might be an indicator of higher stress and cognitive workload (Henderson et al. 1999; Jacob & Karn 2003). This contention is supported by the results of the brain activity analysis of the participants. This analysis shows that, in general, the cognitive load of experts was generally lower than that of novices. In addition, our brain activity analysis indicated that novices, in contrast to experts, had a greater cognitive load using continuous animations than semi-static animations. This is a further confirmation that novices have significantly more difficulties in solving the detection task with continuous animations than with semi-static animations.

Moreover, as discussed in previous Section *5.5.2, Animation Design*, novices with semi-static animations performed the task significantly better than novices with continuous animations, because of the discrete and abrupt change of the dynamic scenes. As aircraft dynamics are not perceived directly, the dynamic contrast among the displayed aircraft, and consequently the pre-attentive effect of the visual cues, became less strong. Semi-static animations could be thus useful to prevent the erroneous detection of perceptually more salient objects, and so facilitate viewers in the identification of thematic relevant visual information.

Our fourth hypothesis (cf. **WH 3.2**) was that the efficiency and effectiveness of task performance with semi-static and continuous animations might be influenced by the *spatial abilities* of users. As Newcombe and Frick (2010) suggest, people can be more successful in a particular academic or professional domain because of their spatial skills. In the domain of air traffic control, operator candidates must possess good spatial skills in order to enter ATC training schools. Their spatial skills will then probably further improve during ATC training and in the course of their everyday job demands. Based on Hegarty et al. (2003), we expected that users might perform the task more effectively and efficiently with the semi-static design type if they have high spatial abilities. Conversely, users might perform the task more effectively and efficiently with the continuous design if they have low spatial abilities.

We found indeed that participants with higher spatial abilities perform better than participants with lower spatial abilities. However, the influence of spatial abilities on task performance might be relevant only for the continuous displays. Supposing that continuous animations require higher processing demands and higher cognitive load compared to semi-static animations, a higher spatial skill level might have a positive influence on completing the task effectively and efficiently. This finding emphasizes the relevance of visuospatial skills and prior knowledge, beyond display design, on task performance. Previous studies of Grabner et al. (2003) show that well-trained people, and/or people that are considered 'more intelligent' (e.g., with higher IQ scores) exhibit lower and more focused cortical-activation, which in turn indicates a higher neuronal efficiency. This means that having superior spatial skills, as well as expertise or prior knowledge, reduces the cognitive effort required to solve the task, and thus positively affects task performance. Further, high spatial skills might be advantageous for people when processing information with novel (i.e., continuous) and even cognitively demanding

animated displays. Training in a specific domain might help people in processing complex information and transferring previous knowledge into unfamiliar and novel domains even without significant compromise of task effectiveness.

According to our fifth hypothesis (cf. **WH 3.3**), we expected that participants' *emotional and affective states*, and consequently their task performance, might be influenced by their expertise level and the animation type used to solve the task. We found indeed that user psycho-physiological responses, self-perceived stress level (i.e., engagement, distress, and worry), and motivation are influenced by their expertise level. Our analysis of participant electrodermal activity (EDA) shows that, in general, experts showed higher arousal levels than novices. We further investigated the participant's self-perceived stress factors (i.e., engagement, distress, and worry) to reliably interpret EDA results. In general, we found that experts showed an increase of engagement and a decrease of distress and worry in completing the experiment compared with novices. This result matches findings in sport research (Saha 2015) and supports the *Inverted-U Hypothesis* (Yerkes & Dodson 1908) in which experts, or elite performers, usually show higher levels of arousal, combined with better performance, than novice performers.

The *Yerkes-Dodson low (Yerkes & Dodson 1908)* states that task performance has a curvilinear (or inverted-U) relationship to the stress level of performers. If participants are too stressed (or less motivated), their performance declines. However, this contention is not supported by our results based on the collected EEG data, which revealed additional information about valence of the participants' affective states. We expected that air traffic controllers would be more motivated than psychology students in participating in the experiment because of the similarity with their job-related tasks, and their familiarity with the displays used in their everyday job (cf. **WH 3.3**). However, according to our EEG analysis, the affective state of novices seemed to be more positive than the affective state of experts, even though their performance tends to be lower. The absolute engagement values of novices was higher than those of experts. This leads us to believe that experts were less motivated than novices to perform the task. We did find a correspondence of our results with past vigilance studies (Matthews et al. 2013). Here, air traffic controllers showed a reduction of engagement, as well. Matthews et al. (2013) argue that reduction in task engagement may be related to monotonous ATC-related tasks requiring continuous attention demands. For experts, these kinds of (monotonous) tasks may reduce energy and

motivation, resulting in fatigue. In contrast, the novelty of the task for novices may have a positive effect on their engagement state.

Furthermore, we hypothesized that the animation type might also influence emotional and affective states of the participants, and consequently their task performance (cf. **WH 3.3**). As mentioned previously, we found that experts show higher physiological arousal (EDA) levels than novices. However, this difference between novices and experts was significant only for the semi-static condition. With continuous animations, EDA levels were generally higher than those with semi-static animations, but this difference was not statistically significant. In addition, the analysis of stress state factors (Helton 2004) showed that the self-assessments of the experts exhibited more engagement, less distress, and less worry with familiar semi-static animations than novices. This pattern was almost mirrored for the continuous animations. Experts and novices showed similarly low engagement, but novices were more distressed and worried than experts. Experts, even if they performed well with both animation design types, showed a decrease in engagement levels, whereas their distress increased with the unfamiliar continuous animations. Conversely, being less distressed and less worried with the semi-static animations has a positive effect on response accuracy for novices with this animation design type. Previous studies (Kriz & Hegarty 2007; Mayer et al. 2005) point out that the temporal constraints and the continuous processing of spatio-temporal information in animations may cause an increase of the cognitive load, which is linked to distress and worry. Similarly, the pauses from one state to another in semi-static animations may reduce cognitive overload and thus have a positive effect on the emotional and affective state of the user (cf. **WH 3.3**). Anecdotal evidence from post-test questionnaires confirms this hypothesis, suggesting that participants found the task more difficult to solve with continuous animations than with semi-static animations.

This chapter can be summarized as follows: with our first experiment in the ATC context, we emphasize the importance of task, animation design, and user characteristics in the apprehension of movement changes with animations. Furthermore, our findings about the affective states of the participants exposed in this chapter highlight the importance of coupling performance measures with other metrics, e.g., participant psycho-physiological responses. This allowed us to gain a deeper insight about participant emotional and cognitive states (e.g., cognitive workload and motivation) during the experiment, and so to reliably interpret our empirical results. We also emphasize the

relevance of top-down mechanisms for experts, contrarily to novices, in information processing with animations. For experts, their familiarity with ATC tasks and ATC displays allowed them to effectively and efficiently perceive and detect task-relevant moving objects, irrespectively of the animation design.

Now, what will happen when experts have to solve the third SA task, i.e., predict future aircraft movements, with different types of animations and task difficulties? This is the question that we attempted to answer in our second empirical study presented in the following chapter, Chapter 6 "*Experiment II*".

CHAPTER SIX

EXPERIMENT II: PREDICTION OF FUTURE MOVEMENT PATTERNS WITH ANIMATIONS

6.1 Motivation and Research Goal

For the first experiment, we designed the test stimuli so that the accelerating aircraft was never the fastest aircraft. On the one hand, this allowed us to examine potential detection differences between thematically relevant information and perceptually more salient information. On the other hand, the perceptually salient objects might have drawn participant attention away from the accelerating aircraft, resulting in lower task performance. The additional question that emerges from the first experiment is how the direct perception of relative motion among moving visual entities with continuous animations and the perceptual saliency of certain movement patterns might positively help experts in the ATC domain in solving specific tasks, and in better understanding what consequences a specific animation design choice could have on the aviation safety. For this reason, for this second experiment, we primarily aim to more systematically assess the way in which *animation design characteristics and relative movements between moving objects influence user perceptions and information extraction of two coordinated spatio-temporal events*.

The results of the previous experiment also reveal that the depiction of aircraft path history is essential for participants to extract motion changes, especially with semi-static animations. With semi-static animations, participants derived speed changes directly from the length and size of the radar comet shape (i.e., the aircraft position history), as movement changes are not perceived directly through the objects' motion. Conversely, with continuous animations, movement changes are encoded twice, i.e., they are perceived directly through the continuous motion of moving objects and though the change of the features' visual form. A further question that arises from this first experiment is *how essential is the depiction of aircraft path history in continuous animations* to effectively detect spatio-temporal information. In other words, this question addresses whether the motion itself, i.e., the depiction of only the current aircraft position without past positions, might be sufficient to efficiently and effectively identify specific aircraft movement patterns.

Moreover, previous studies with semi-static ATC displays demonstrated that *aircraft speed and temporal proximity/distance of minimum separation between two converging aircraft (i.e., distance between two aircraft at the point of closest passing)* have a significant influence on air traffic controller performance relative to conflict detection (Rantanen & Nunes 2005; Boag et al. 2006). In general, air traffic controller performance decreases by increasing the relative speed difference and/or the distance of minimum separation between aircraft. It is not clear if the same principles are also valid for continuous animations (Schlienger et al. 2007), and this is the third question that we wish to address for this second user study.

In general, the understanding of aircraft dynamics and the anticipation of future actions is an important mechanism in a situation awareness (SA) context. Neisser (1967) points at the importance of *anticipation skills* as precursors of human decision-making processes, such as in sports or surveillance domains with real-time movement data. Focusing and correctly perceiving future states of spatio-temporal events or phenomena increase task performance and support users positively in their decision-making processes, such as in the air traffic control domain (Durso et al. 2006). However, superior performance in predictive tasks can be achieved not only by means of adequate perceptual, affective, and cognitive capabilities, but also by means of an optimal interaction with an appropriately designed animation.

For this experiment, we identified three compound movement patterns that are particularly relevant for motion tracking in ATC: parallel, convergent, and divergent

aircraft patterns (Dodge 2015). Parallel and divergent patterns are not associated with specific critical situations, whereas convergent patterns are associated with the occurrence of potential critical situations for air traffic security. Tracking convergent aircraft movements are thus particularly relevant in ATC, because it presumes the prompt recognition of air traffic conflicts. For this reason, we devoted particular attention to convergent movement patterns as a real-world motion-tracking task (compared to other kinds of movement patterns) when testing animated displays.

For this second experiment, we thus investigated how *the context of use and task* (i.e., real-world SA-task in the ATC context), *animation design characteristics* (i.e., rate of change and depiction of aircraft past positions), and *user-related factors* (i.e., air traffic controllers' training level and affective state) might affect participant task performances (i.e., response accuracy) in anticipating future aircraft movements. Moreover, we also aimed to better understand and more systematically assess how users conceive of and predict coordinated spatio-temporal events with animations, how the effect of relative motion among moving objects presented on different animated display types might influence user decision-making processes, and what consequences this might have for air traffic security.

6.2 Specific Research Questions and Working Hypotheses

For this second experiment, we tested the following research questions relative to the context of use and task, the animation design, and user-related factors:

- **SRQ 1 (CONTEXT OF USE and TASK):** How does *task difficulty* (i.e., *different relative speed* and *distance of minimum separation between two converging aircraft*) influence the prediction of future aircraft movements in animations?
- **SRQ 2 (ANIMATION DESIGN):** How does the *animation design type* (i.e., *semi-static vs continuous animations* and *displays depicting path history vs displays without path history*) influence the prediction of future aircraft movements in animations?
- **SRQ 3 (USER):** How do specific *user-related factors* (i.e., perceptual, cognitive, and affective processes, including individual and group differences) influence the prediction of future aircraft movements in animations?
    - ***SRQ 3.1:*** Is task performance correlated with the *electrodermal activity, training, and affective states of the users, as well as with task difficulty*?
    - ***SRQ 3.2:*** Does the *training level* with semi-static displays influence *user preferences, ease of use, and EDA values*?

Our ***working hypotheses*** concerning the above-mentioned specific research questions were as follows:

(1) Task performance and cognitive load would increase by increasing *task difficulty*, i.e., by presenting animations displaying aircrafts with different relative speeds and with increasing minimum separation distances (Rantanen & Nunes 2005; Boag et al. 2006). The visual processing of displays depicting aircrafts moving at different speeds would be more difficult and require higher working memory capabilities than with aircrafts moving at the same speed. The same principle would be valid for increasing minimum separation distances between aircraft. The farther two aircraft are positioned from each other, the more difficult it would be for users to predict future aircraft states.

WH 1 (CONTEXT OF USE and TASK): *Task performance decreases with an increase of the task difficulty, i.e., by presenting animations displaying aircrafts with different relative speeds and with increasing minimum separation distances.*

(2) Task performance would be greater with *continuous animations* than with semi-static animations, because continuous animations facilitate users creating a mental model of aircraft movement dynamics. Consequently, this allows them to effectively perceive relative motion and to effectively predict future aircraft positions. However, air traffic controllers are trained in using semi-static animations and not continuous animations in their everyday work. This training with this display design might influence task performance, as well. In addition, task performance would increase, and cognitive load decrease, with the inclusion of past positions (*path history or trace*) in the visualization of aircraft movements.

WH 2 (ANIMATION DESIGN): *Continuous animations and path history improve task performance by helping controllers to easily perceive relative motion and thus to easily predict future aircraft patterns (compared to semi-static animations and displays without path history).*

(3) We expected that user-related factors, i.e., their affective state during the experiment, would affect their task performance. Their affective state would be, in turn, influenced by their training level with the animated displays, and by the task difficulty. We expected that controllers with less training would perform less well, and show a more negative affective

state (i.e., more distress and worry, less engagement) and higher EDA responses, compared with well-trained controllers. In addition, we supposed that controllers, because of their training and familiarity with these display types, would perform better with standard semi-static animations and with displays showing path history. Since positive emotions (e.g., pleasure) seem to be correlated with EDA values and familiarity (van den Bosch et al. 2013), we expected that controllers would show higher EDA values with the above-mentioned standard animation types. This correlation would also be positively reflected in their self-reported design preferences and easiness scores.

WH 3.1 (USER): *Task performance increases with a decrease of stress factors (i.e., more distress, more worried and less engagement), and respectively of EDA responses and task difficulty, and with an increase of training.*

WH 3.2 (USER): *Training and familiarity with semi-static displays, and with displays showing path history, influence positively controllers' EDA values, design preferences and ease of use ratings. We (and they) will expect to perform better with semi-static displays depicting path history than with the other animation design types.*

6.3 Experimental Design

We designed a mixed factorial experiment according to a within-subject design (i.e., (2 x 3) x (2 x 2) test factors). We manipulated the following two independent variables (IV):

- ***IV 1 (CONTEXT OF USE and TASK):*** We manipulated two factors: (F1) the relative speeds, and (F2) the minimum separation distances, between two converging aircraft.
  The first factor, i.e., the **relative speed between two converging aircraft (F1)**, had two levels of difficulty:
  - **Difficulty level 1 (ES):** Two aircraft moving at equal speed (i.e., both aircraft moving at 250 kts).
  - **Difficulty level 1 (DS):** Two aircraft moving at different speeds (i.e., one aircraft moving at 160 kts and one at 250 kts).

  The second factor, i.e., the **minimum separation distance between two converging aircraft (F2)**, had three levels of difficulty:
  - **Difficulty level 1 (0 Nm):** Minimum separation distance of 0 nautical miles (Nm) between the two aircraft. This corresponded to a *collision situation*, respectively to the most critical situation in air traffic control.
  - **Difficulty level 2 (3 Nm):** Minimum separation distance of 3 Nm between the two aircraft. This corresponded to the minimum safe distance that two aircraft have to respect on the horizontal plane. A distance smaller than 3 Nm corresponds to a *conflict situation*. Air traffic controllers are especially trained to promptly recognize minimum separation distances between 0 and 3 Nm.
  - **Difficulty level 3 (6 Nm):** Minimum separation distance of 6 Nm between the two aircraft. This corresponds to a *no conflict situation*.

- ***IV 2 (ANIMATION DESIGN):*** We manipulated two factors: (1) the *rate of change* of the animated displays (F1), and (2) the *depiction of path history* (F2).
  The first factor, i.e., the **rate of change of the animated displays (F1)**, had two levels:
  - **Level 1 (S):** Semi-static animations (standard animation design type that air traffic controllers use in their everyday jobs).
  - **Level 2 (C):** Continuous animations (novel animation design type for air traffic controllers).

  The second factor, i.e., the **depiction of path history (F2)**, had two levels:
  - **Level 1 (P):** Animations depicting path history (i.e., aircraft current position, as well as its past positions, are depicted).
  - **Level 2 (nP):** Animations without past positions (i.e., only aircraft current position is depicted).

The choice of the above-mentioned independent variables and the development of the experimental design were the result of several discussions with ENAC and of a pilot test conducted with air traffic controllers at ENAC. To summarize, Table 3 illustrates the independent variables used for this experiment, including their corresponding factors and levels.

Table 4: *Summary of the independent variables for the Experiment II with their corresponding factors and levels.*

INDEPENDENT VARIABLE	FACTOR	LEVEL
CONTEXT OF USE and TASK	*Relative speed*	Aircraft moving at equal speed
		Aircraft moving at different speeds
	Minimum separation distance	0 Nm (collision)
		3 Nm (minimum safe distance)
		6 Nm (no conflict)
ANIMATION DESIGN	*Rate of change (or smoothness of the transitions between scenes)*	Semi-static animations (1 display every 4 seconds); abrupt transition between scenes
		Continuous animations (60 displays every second; 60 Hz); continuous transitions between scenes
	Depiction of path history	With path history
		Without path history

6.3.1 **Participants**

Twelve air traffic controllers working at Brno Airport in the Czech Republic took part in the experiment, 2 females and 10 males. On average, they were 38 years old and had 8 to 10 years of training and working experience in ATC. We divided the participants into two groups according to their ATC expertise level. As shown in Figure 37, six participants had less than 10 years of ATC expertise (i.e., participants with less training in ATC), and five participants had more than 10 years of ATC expertise (i.e., participants with more training in ATC).

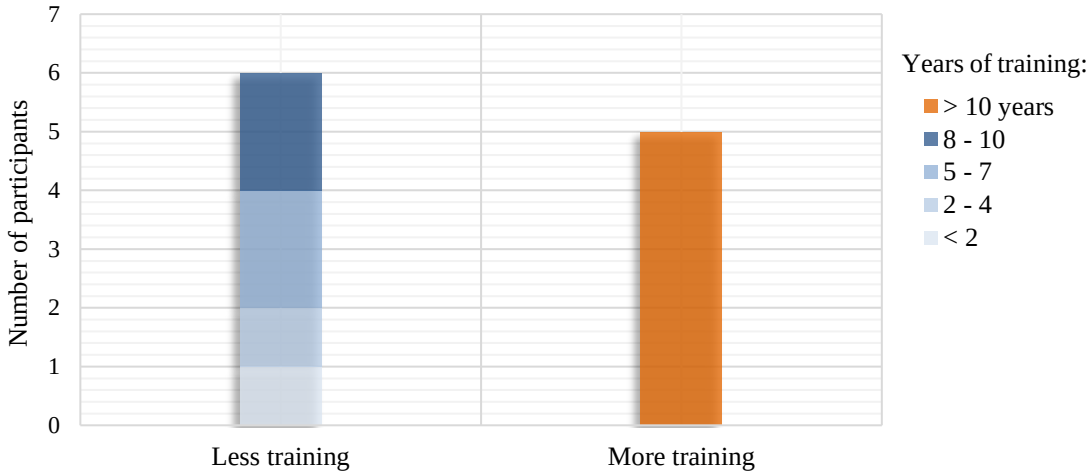


Figure 37: *Participant training levels in ATC (in years). Participants were divided into two groups: less training (i.e., < 10 years) and more training (>= 10 years).*

6.3.2 Materials

Questionnaires for Testing Usability Metrics

To investigate how well participants performed with familiar and novel design types, we prepared a questionnaire according to Albert and Dixon (2003), to be completed prior to and following the experiment. The questions before and after the experiment are provided in

Annex *5: Introduction and Questionnaire for Testing Usability Metrics of Experiment* ***II*** and Annex 7: *Post-Test Questionnaire of Experiment II*.

Since participants were trained, and they worked daily with semi-static animations depicting path history, we expected that they would judge their performance as being better with this kind of displays compared to continuous animations, and displays without path history. After the experiment, we also asked them to assign ease of use rating scores, on a Likert 5-point scale, for the four animation design types.

Test Stimuli

As illustrated at the beginning of this section, we developed a factorial experiment following a within-subject design. In all, participants processed 24 animations (i.e., (2 x 3) x (2 x 2)=24 test stimuli). The 24 animations were designed according to French *Operational Display System* (ODS) radar screens (for more information about ODS displays, see Sections 3.2.2, *Animated ATC Radar Displays*, and 5.3.2, *Materials*).

The 24 test stimuli were shown, in a random order, to the participants, according to a combination of the animation design types mentioned above (i.e., combination of the factors "rate of change" and "path history"). We thus created the 24 displays according to the following animation design types: *six semi-static animations with path history (SP)*, *six semi-static animations without path history (SnP), six continuous animations with path history (CP),* and *six animations without path history (CnP)*. Each display lasts 20 seconds and depicts two converging aircraft always moving with the same heading and in the same direction (from left to right on the display), at a 60° angle to one another. Figure 38 shows two test stimuli with two converging aircraft (i.e., A1 and A2) moving at equal or different speeds. Depending on their initial speeds and positions, if they continue moving at the same speed and heading, they will be closest to one another at a specific minimum separation distance in space. This specific location, i.e., the minimum separation distance between A1 and A2, is represented in Figure 38 with an orange line. The two orange points at the far ends of the line correspond to the future position of the two aircraft when they will reach that distance. However, participants did not see this line, because it is what they were asked to estimate during the experiment. Section 6.3.4, *Procedure*, explains in more details the task and procedure of the experiment.

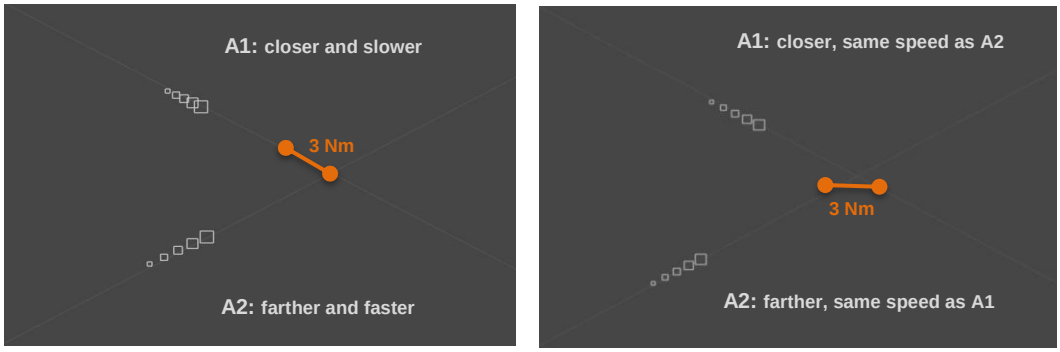


Figure 38: *Example of two test stimuli in DS (on the left) and ES (on the right) condition with 3 Nm of minimum separation between the two aircraft A1 and A2.*

The six displays in each animation design type (i.e., SP, CP, SnP, and CP) were additionally manipulated according to the two task difficulty levels mentioned above (i.e., the relative speed and minimum separation distances between the two converging aircraft). In three displays the aircraft moved at the same relative speed of 250 kts (i.e., *ES condition*), and in three displays, aircraft moved at two different relative speeds, at 160 kts and at 250 kts (i.e., *DS condition*). The test stimuli in each condition (i.e., ES and DS) showed the two converging aircraft at three different minimum separation distances, i.e., 0 Nm, 3 Nm, and 6 Nm. One aircraft (i.e., A1 in Figure 38) was either closer to the potential crossing point (in the case where the two aircraft moved at the same speed) and/or was moving slower. The other aircraft (i.e., A2 in Figure 38) was farther from the potential crossing point (in the case where the two aircraft moved at the same speed) and/or the faster one. For example, Figure 38, left depicts A1 and A2 with different relative speeds, i.e., A1 moves slower than A2. Conversely, Figure 38, right depicts A1 and A2 moving at equal speed, but A1 is closer than A2 to the potential crossing point. In both cases, the minimum separation distance between the two converging aircraft is 3 Nm.

6.3.3 Collected Data and Test Setup

For this second user study, we recorded the following data as dependent variables: the x,y screen coordinates of the two estimated aircraft positions, and the estimated distance between these two aircraft positions (i.e., accuracy of responses), the task completion time electrodermal activity with the e4 wristband,[18] and eye movements tracked with a Tobii

[18] E4 wristband: https://www.empatica.com/e4-wristband/.

TX300 eye tracker (Figure 39).[19] In addition, we collected participant judgements about the ease and preference of the four animation design types (i.e., SP, CP, SnP, and CnP) for solving the required task with a questionnaire. We also measured their self-reported affective state with the Short Stress State Questionnaire (SSSQ) (Helton 2004) and obtained their background information (i.e., age, gender, and training level in ATC) by means of a pre-test questionnaire (see Annex 6: *Pre-Test Questionnaire of Experiment II*).

Figure 39: *Test setup with the Tobii eye tracker and e4 wristband.*

6.3.4 **Procedure**

Prior to running the experiment, participants were asked to complete a questionnaire about their performance expectations in solving the required task (see Annex 5: *Introduction and Questionnaire for Testing Usability Metrics of Experiment II*), followed by a background questionnaire (see Annex 6: *Pre-Test Questionnaire of Experiment II* and the pre-test SSSQ questionnaire (Helton 2004) (see Annex 1: *Short Stress State Questionnaire*). After calibrating the eye-tracker and having completed a short training session with three animations, participants were asked to process the 24 test stimuli on a screen with a resolution of 1680 x 1050.

The *task* assigned to the participants was to observe the dynamics of the two moving aircraft and to estimate their future positions and distances when they would be closest to one another on the display, i.e., by their lower minimum separation distance, given the same initial heading and speed. They watched the animations for 20 seconds and subsequently predicted the future positions of, and distances between, the two converging aircraft by

[19] Tobii TX300 eye tracker: http://www.tobii.com/.

drawing two points connected by a line on the screen with two mouse clicks. Figure 40 shows an example of a test stimulus depicting two converging aircraft moving at equal speed. It indicates as well the response of one participant at the required task with a red line connected by two points. The true minimum separation distance between the two aircraft, as well as their true future positions when they are closest to one another, is highlighted with a green line, respectively with two green points.

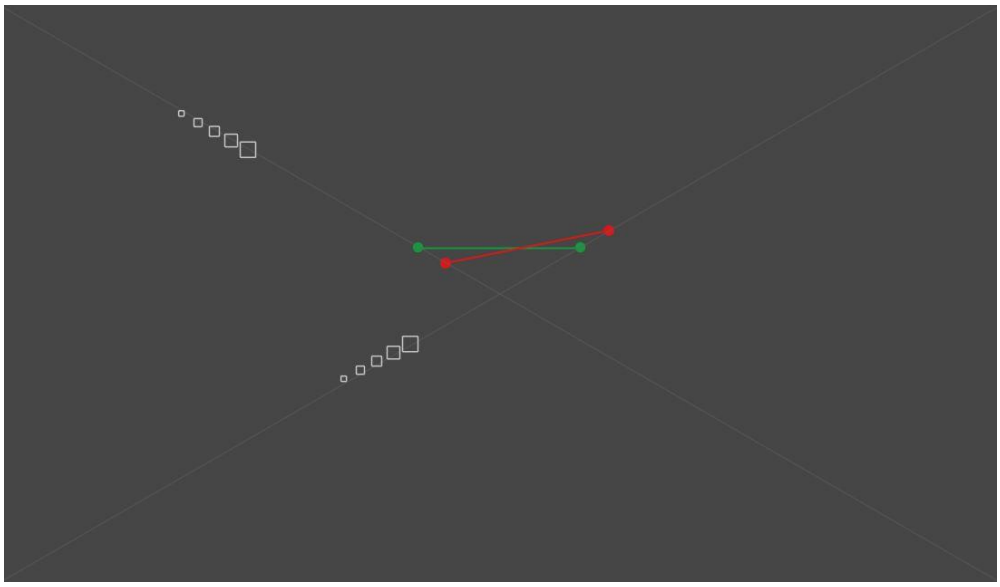


Figure 40: *Example of a test stimulus depicting two converging aircraft moving at the same speed, and showing the response of a participant (red line), as well as the true answer (green line), to the given task.*

After the experiment, participants were asked to fill in a post-test questionnaire with questions about their expectations of and preferences after using the different animation design types (see Annex 7: *Post-Test Questionnaire of Experiment II*), and, again, the SSSQ questionnaire. The entire experimental duration was about 30 minutes: the animation portion, including eye calibration with Tobii studio, took about 20 minutes, and about 10 minutes was dedicated to the questionnaires.

6.4 Results

An important task for air traffic controllers is to correctly predict minimum separation distances between aircraft, and thus to accurately perceive their relative motion to prevent air traffic conflicts. To better understand how the controllers perceived the relative motion between two converging aircraft using the different animation designs, we analysed the results from the questionnaires and their response accuracies.

In this section, I first present results of the pre- and post-test questionnaires, by showing participants' expectation ratings, easiness, and preferences of using the different animation design types. Then, I present participants' response accuracies, in particular, how animation design and task difficulty influenced their estimations of minimum separation distance. Finally, I report results of the participants' SSSQ scores and their EDA responses by comparing them with their task performance.

6.4.1 Questionnaire Responses

Expectation and Experience Ratings

Contrarily to our hypothesis (cf. **WH 3.2**), questionnaire responses revealed that participants expected, before starting the experiment, to perform the estimation task more accurately with continuous animations depicting path history (CP), followed by the standard semi-static animations depicting path history (SP), and then by both animated displays without path history (CnP and SnP). Similarly, after the experiment, they rated their task performance as more accurate with CP, followed by SP, CnP, and SnP. In addition, they judged their task performance with SnP and CnP as lower than their initial expectations (Figure 41).

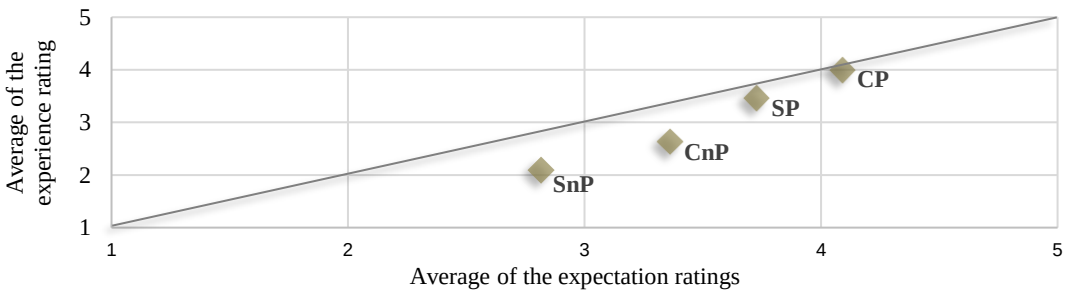


Figure 41: *Expectation and experience ratings (according to methodology of Albert and Dixon (2003)) across the four display design types.*

Ease of Use with the Different Design Types and Design Preferences

As shown in Figure 42, participants found it easier to solve the task with the continuous animations depicting path history (CP) (M=3.91, SD=0.83), compared to the other three design types. The second-easiest design type was SP (M=3.36, SD=1.03), followed by the CnP (M=2.36, SD=0.81). The SnP (M=1.73, SD=0.79) was judged as the most difficult design type to use to estimate future aircraft positions (Figure 42). Rating scores differed significantly across the four design types ($\chi^2(3)=23.82$, $N=11$, $p<.000$). However, the rating

scores between CP and SP were not significantly different. This result also differs from our hypothesis (cf. **WH 3.2**).

As a consequence, the participants' preference rating scores about the animation design type resulted in similar values as their ease-of-use scores. As shown in Figure 43, participants preferred to solve the task mostly with CP, followed by SP, CnP, and SnP, as interface design types to solve the task. As for their ease-of-use scores, their preference ratings between CP and SP are not significantly different.

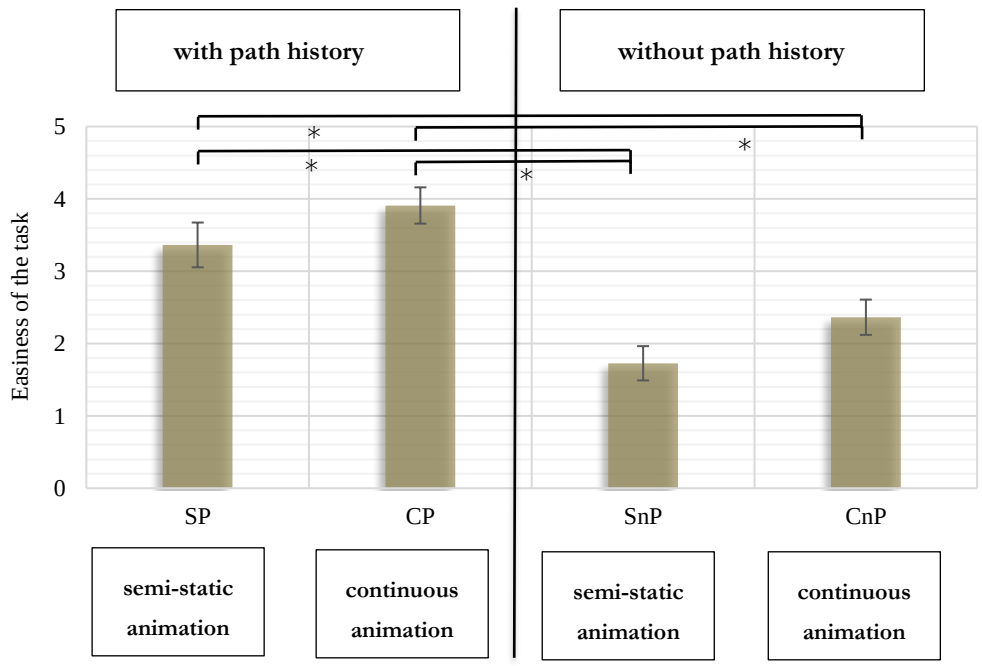


Figure 42: *Mean rating scores about task easiness across the four animation design types on a 5-point Likert scale (1="very difficult", 5="very easy").*

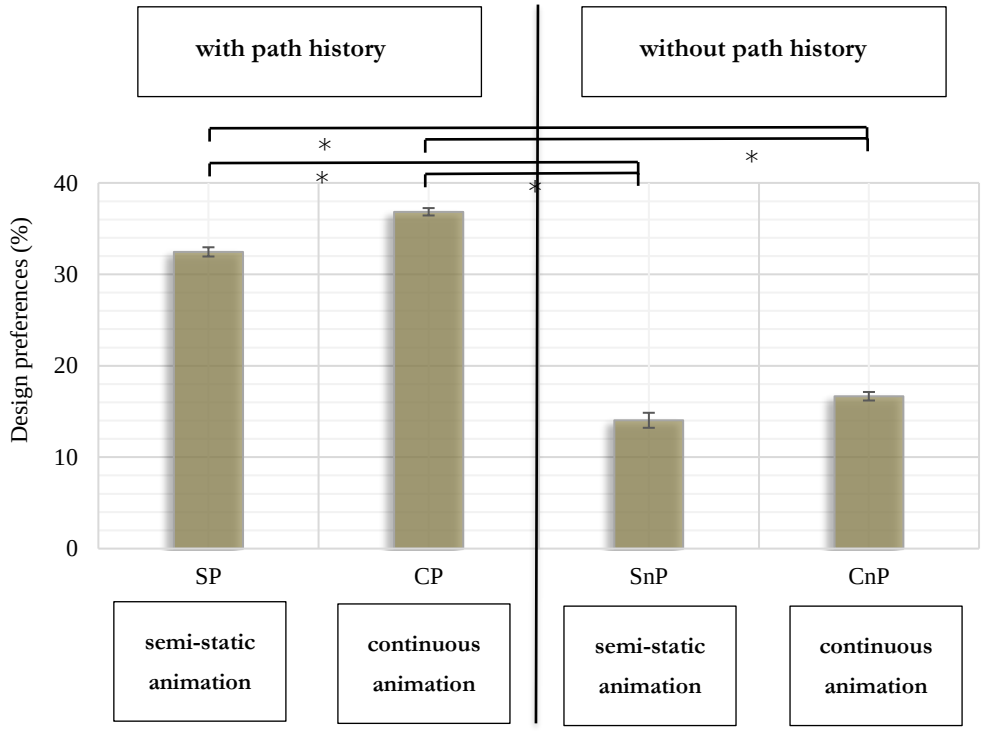


Figure 43: *Mean ranking scores of participants' design preferences (%).*

Open Questions

Participants had the opportunity to explain the reasons for their preferences between semi-static and continuous animations. They reported that continuous animations allow the acquisition of more precise information about the dynamics and relative speeds of aircraft. In their opinion, this may facilitate their predictions about future positions and movement dynamics of aircraft. Regarding the depiction of aircraft path history, participants judged it to be very important, and mostly essential, to understand current movement dynamics and predict future movement behaviour of aircraft, in both semi-static and continuous conditions.

In the following section, participant response accuracies are provided, which helped to better understand how controllers actually estimated the *relative motion* between two converging aircraft with different animation designs.

6.4.2 Quantitative Analysis of the Participants' Response Accuracies

We firstly calculated participant estimation accuracies (%) relative to the minimum separation distances between two moving aircraft across the four animation design types (i.e., SP, CP, SnP and CnP). We calculated the normalized distance errors with Formula 1. The variable $ERROR$ corresponds to the difference between the estimate and the correct value of the minimum separation distance between the two converging aircraft. The variables $ERROR_{max}$ and $ERROR_{min}$ refer to the maximum and minimum error of each participant's sample data, respectively.

$$RAD_{perc} = \frac{ERROR}{ERROR_{max} - ERROR_{min}} * 100$$

Formula 1: *Response accuracy of distance estimation in %.*

Overall, we found that participants estimated distances between the two aircraft differently across the four animation design types ($F(2.06, 22.68) = 5.67$, $p<.010$). On average, they estimated distances most accurately with continuous animations with path history (CP), followed by semi-static animations with path history (SP), continuous animations without path history (CnP), and, the least accurately assessed display, semi-static animations without path history (SnP) (Figure 44).

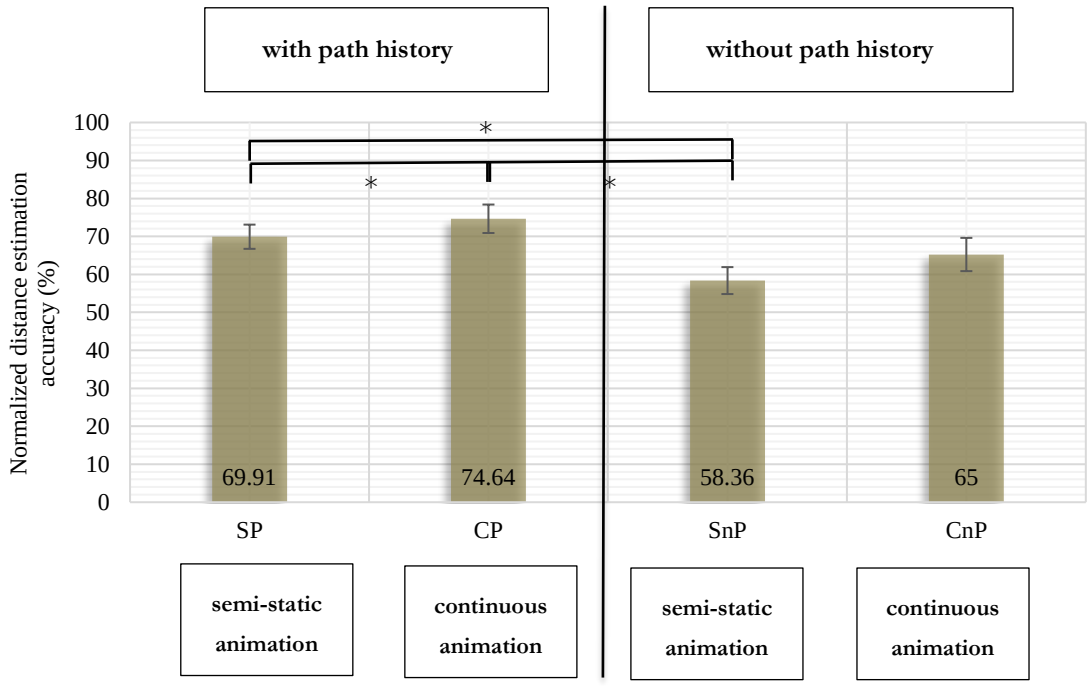


Figure 44: *Normalized distance estimation accuracy (%) across animation design types.*

Significant differences were found between SP and CP (p<.032), between SP and SnP (p<.014), and between CP and SnP (p<.001). As hypothesised (cf. **WH 2**), we thus found an effect of the rate of change and path history, with which aircraft movements are depicted, on participant estimation accuracies of minimum separation distances.

We further analysed the influence of task difficulty levels on participant estimation accuracies. We compared participant response accuracies between the three displays with aircraft moving at equal speed (i.e., ES condition) and the three displays with the aircraft moving at different speeds (i.e., DS condition). As hypothesized (cf. **WH 1**), participants performed the task more accurately in the ES condition than in the DS condition (F(1,11)=7.37, p<.020) (Figure 45).

Across the four animation design types, participants responded more accurately with the continuous animations with path history (CP) and less accurately with the semi-static animations with path history (SP), followed by the continuous animation without path history (CnP) and then the semi-static animations without path history (SnP). In the DS condition, participants estimated distances between aircraft significantly more accurately with CP than with SnP (F(1,11)=20.67, p<.001). In the ES condition, no significant differences across the four animation design types were found.

The above-mentioned results show that the smaller the difference of participant estimation accuracies is, the better they perceive the relative motion between the two converging aircraft. As hypothesised (cf. **WH 2**), the smallest difference, and thus the more accurate perception of the relative motion between moving aircraft, was observed with the continuous animations depicting aircraft with path history (CP), compared to the other animation design types. In addition, we found that speed differences between aircraft also influenced participants' estimation accuracies (cf. **WH 1**). In general, with displays depicting aircraft at equal speed, participants responded more accurately than with displays showing aircraft at different speeds. In this latter case, participants perceived one aircraft as being faster or slower than the real speed.

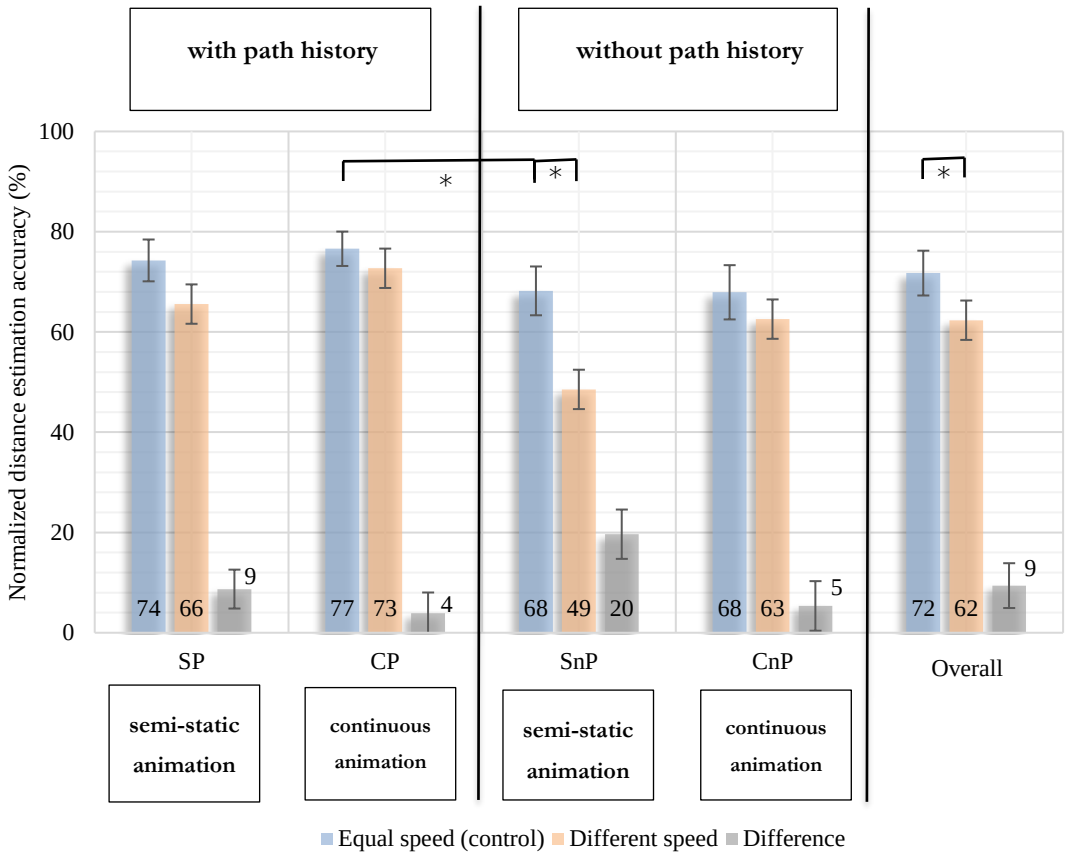


Figure 45: *Normalized distance estimation accuracy (%) between displays depicting aircraft with the same speed (ES) and displays depicting aircraft with different speeds (DS), and between the four animation design types.*

As for the position estimations, we found a significant effect of the minimum separation distance on participants' response accuracies. On average, there is a significant difference of the estimation accuracies across the three tested minimum separation distances (i.e., 0 Nm, 3 Nm, 6 Nm) ($F(19.00,1.73)=25.18$, $p<.000$), as shown in Figure 46. However,

participant task performance decreases significantly only with an increase of the minimum separation from 0 Nm to 6 Nm (p<.000), and from 3 Nm to 6 Nm (p<.001), but not from 0 Nm to 3 Nm. Consequently, participants estimated distances between the two aircraft in a more accurate way, when aircraft are in a potentially critical situation, i.e., when they are separated by a distance between 0 Nm and 3 Nm to one another, compared to a situation without conflicts (i.e., aircraft separated by a distance of 6 Nm). However, this finding is not valid for the continuous animations with path history (CP), in which we did not find any significant effect of the minimum separation distance between aircraft on participant task performance.

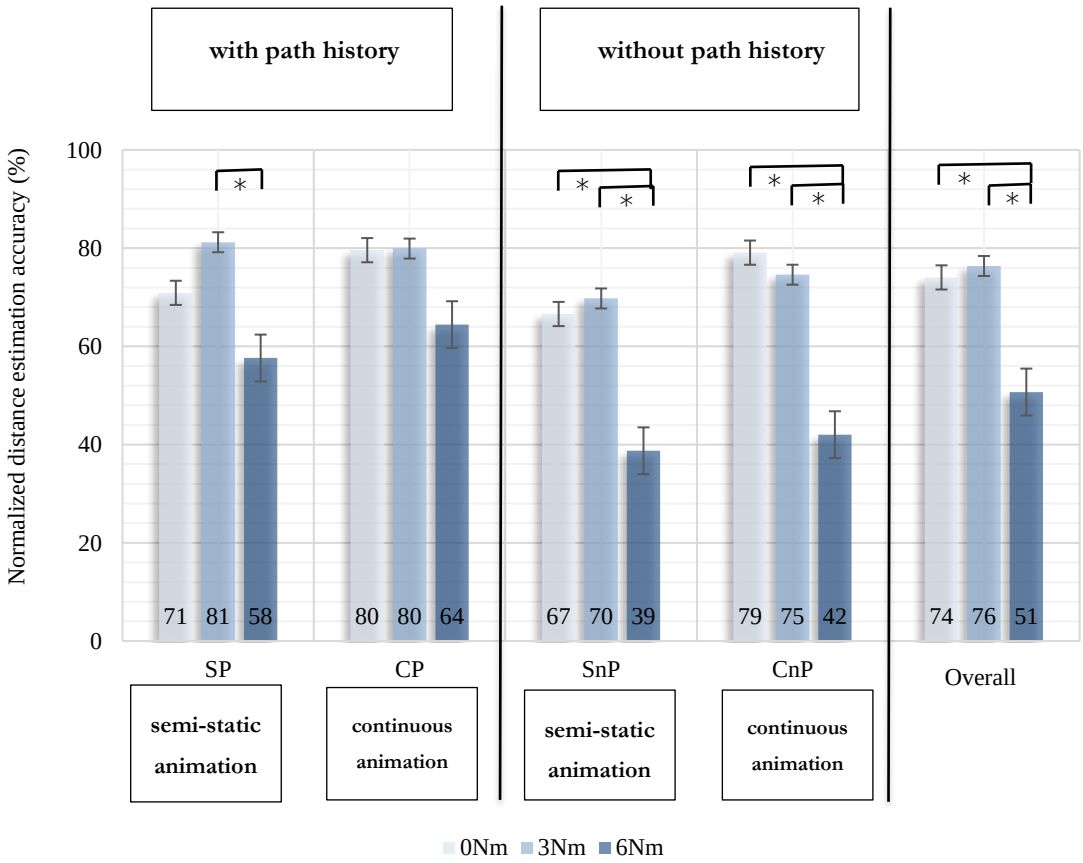


Figure 46: *Distance estimation accuracy according to three distances of minimum separation (i.e., 0 Nm, 3 Nm, and 6 Nm) between aircraft.*

6.4.3 **Qualitative Analysis of Participant Response Accuracies**

To better explain why participant response accuracies differed across the four animation design types, and to better understand how they perceived the relative motion between the depicted moving objects, we further assessed their distance estimations qualitatively. In addition, we investigated the potential consequences that different animation design types might have in the air traffic domain for particular safely critical tasks.

More specifically, we qualitatively examined participant position estimations in relationship with the true positions of the two converging aircraft (i.e., A1 and A2) when they were closest to one another. To do that, we first calculated the *error difference* between the estimated aircraft positions and the true aircraft positions. Successively, we calculated the error difference between the two estimated aircraft positions with the formula $\Delta E = abs(E_2 - E_1)$, in which E_1 corresponds to the estimation error of the first aircraft position (A1) and E_2 the estimation error of the second aircraft position (A2). This means that the smaller the error difference (i.e., ΔE) is, the more accurately a participant estimated the future positions, and thus the relative motion and minimum separation distance, of the two converging aircraft.

However, both position estimation errors and the *perception of the depicted aircraft speeds* are important to better understand how participants predicted minimum separation distances between the two aircraft. The way in which participants perceived aircraft speeds had a direct consequence on the estimation of the minimum separation distance as well. For example, if the future aircraft positions had been estimated before the true positions, this means that the aircraft motion was perceived to be *slower* than the true one. If future positions were estimated beyond true positions, this meant that the aircraft motion was perceived to be *faster* than the true one. Consequently, if both aircraft are perceived to be faster or slower than true speeds, then the estimation of the minimum separation distance was more accurate compared to the case of one aircraft perceived as slower and the other aircraft faster than the true speeds.

In our analysis, we further assigned a positive or negative value to the perceived aircraft speeds. If aircraft speeds were perceived to be slower than true speeds, then their values were assigned a negative score. In contrast, if aircraft speeds were perceived to be faster than true speeds, then their values were assigned a positive score. Successively, we calculated the difference between the scores of the two aircraft speeds, across all

participants and test stimuli, according to the formula $\Delta V = abs(V_2 - V_1)$, with V_1 for the A1 scores and V_2 for the A2 scores. If the two aircraft positions were estimated with a similar positive (or negative) value, than ΔV showed low scores. In contrast, if one aircraft was perceived to be slower and the other aircraft faster than true speeds, then ΔV showed greater values. As shown in Table 5, both factors, i.e., the error difference between the two aircraft positions (ΔE) and the perception of aircraft speeds (i.e., V1 and V2), gave us a deeper insight about how participants perceived the relative motion between the two aircraft, and thus, how they accurately estimated the minimum separation distance between the two aircraft.

Table 5: *The combination of estimation error of aircraft future positions and perception of aircraft speeds gave us a better understanding on how participants perceived the relative motion, and thus accurately estimated the minimum separation distance, between two converging aircraft (moving at equal or different speeds).*

		Estimation errors of aircraft future positions	
		E1 and E1 are small	**E1 and E2 are great**
Perception of aircraft speeds (ΔV)	**Both aircraft have positive or negative scores (ΔV is small)**	The relative motion between aircraft is accurately perceived and the minimum separation distance between aircraft accurately estimated; the speed of both aircraft has been accurately perceived	The relative motion between aircraft is accurately perceived and the minimum separation distance between aircraft accurately estimated, but both aircraft is perceived to be much slower or faster than true speeds
	One aircraft has a positive score and one aircraft has a negative score (ΔV is great)	The relative motion between aircraft is not accurately perceived and the minimum separation distance between aircraft is not accurately estimated; the speed of one aircraft is perceived to be slower, and the speed of the other aircraft to be faster, than the true speeds	The relative motion between aircraft is not accurately perceived and the minimum separation distance between aircraft is not accurately estimated; the speed of one aircraft is perceived to be much slower, and the speed of the other aircraft to be much faster, than the true speeds

For example, as shown in Figure 47, a participant perceived both aircraft (i.e., A1 and A2) speeds to be faster than true speeds, and their position estimation errors were relatively low (i.e., E1 and E2, with respect to the other participants' estimations) for both aircraft. In this case, the V-scores of the two aircraft (i.e., V1 and V2) were thus both positive and ΔE was low. This means that the estimated minimum separation distance between the two aircraft (red line in Figure 47) was similar to the true one (green line in Figure 47).

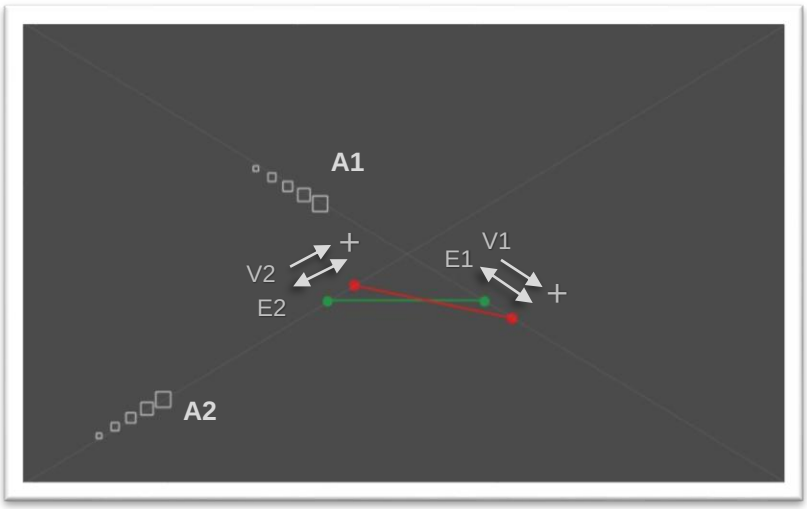


Figure 47: *Example of a participant's estimation of the future aircraft positions (the red points), in which ΔE is low and both aircraft speeds were perceived to be faster than the true speeds (their V-scores are both positive). In this case, the prediction of the minimum separation distance between the two aircraft was accurate (i.e., the true distance depicted as green line is similar to the estimated distance depicted as red line).*

In contrast, if ΔE and ΔV were larger, participants apprehended the motion of the two aircraft differently. For example, as shown in Figure 47, if a participant estimated the future position of the faster aircraft (i.e., A2) well, but with a lower accuracy for the slower aircraft (i.e., A1), e.g., perceiving that it moved faster than in reality, it meant that the participant erroneously perceived the two aircraft as moving at similar speeds. In this case, the estimated minimum separation distance is shorter than the true distance.

The deviations of participant estimations of the minimum separation distance (i.e., if they are shorter or longer) from the true values can be seen in Figure 48. In this figure, the higher the depicted values (%), the shorter the minimum separation distance between the two aircraft estimated. Values equal to 50% mean that estimated distances correspond to the true ones, whereas values below 50% mean that estimated distances are greater than the true ones.

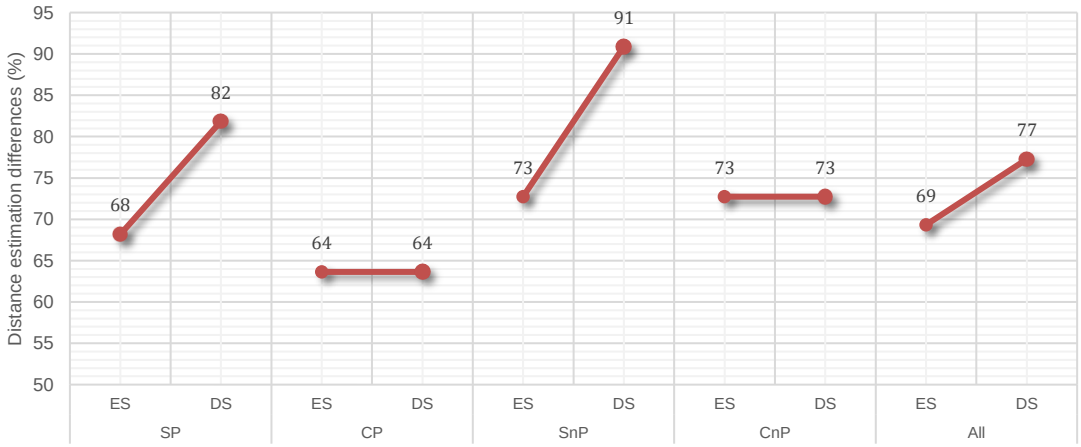


Figure 48: *Differences in the participants' minimum separation distance estimations (%) between estimated values and true values, and across animation design types and difficulty levels (i.e., ES and DS). The higher the depicted value (%) is, the shorter the minimum separation distance has been estimated.*

Figures 49, 50 and 51 show participant's position estimation errors (i.e., ΔE) and their perception of the depicted aircraft speeds across animation design types (i.e., SP, CP, SnP, and CnP) and task difficulty (i.e., ES and DS). Dotted lines show the scores of participant aircraft speed perceptions (light blue for the slower or/and closer aircraft A1, and dark blue for the faster or/and farther aircraft A2). Speed scores between 50% and 100% are positive, while scores below 50% are negative. The higher the positive scores, the faster participants perceived the aircraft speed. Further, the green line shows differences between the perception of the two aircraft speeds (ΔV). The smaller this difference, the more accurate the estimation of relative motion between the two aircraft. Finally, the violet line shows differences in the position estimation errors between the two aircraft (ΔE).

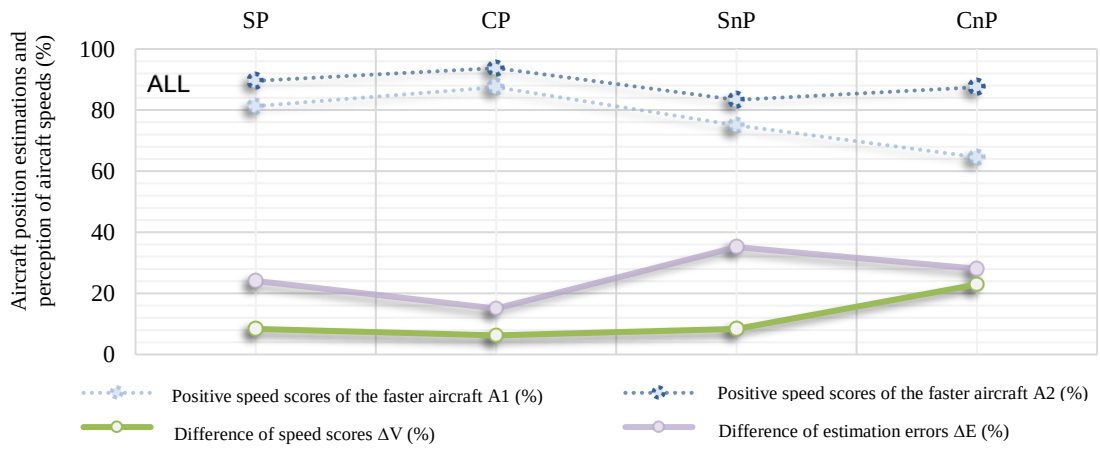


Figure 49: *Position and distance estimation differences between the two aircraft A1 and A2 for all stimuli.*

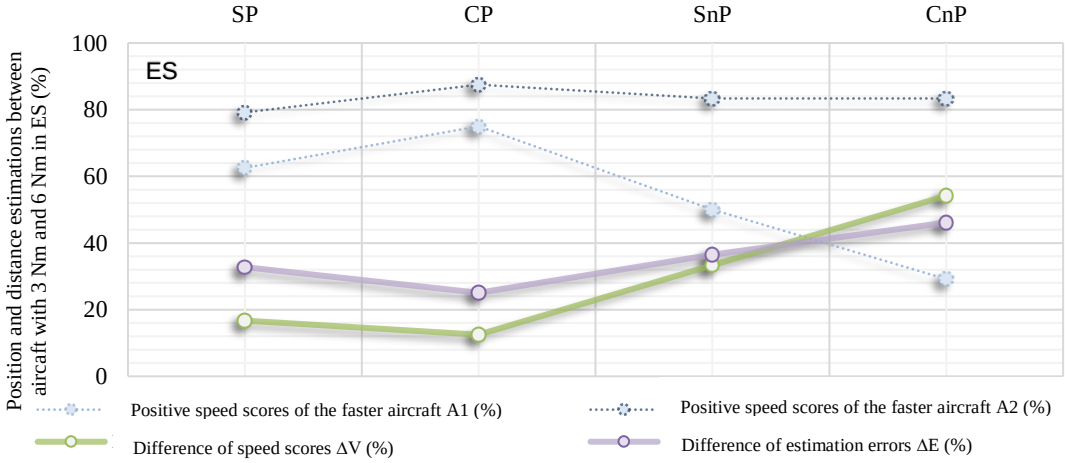


Figure 50: *Position and distance estimation errors of the two aircraft (i.e., closer aircraft A1 in light blue/red and farther aircraft A2 in dark blue/red) relative to the correct future positions for all stimuli in the ES condition and with 3 Nm and 6 Nm distance separation.*

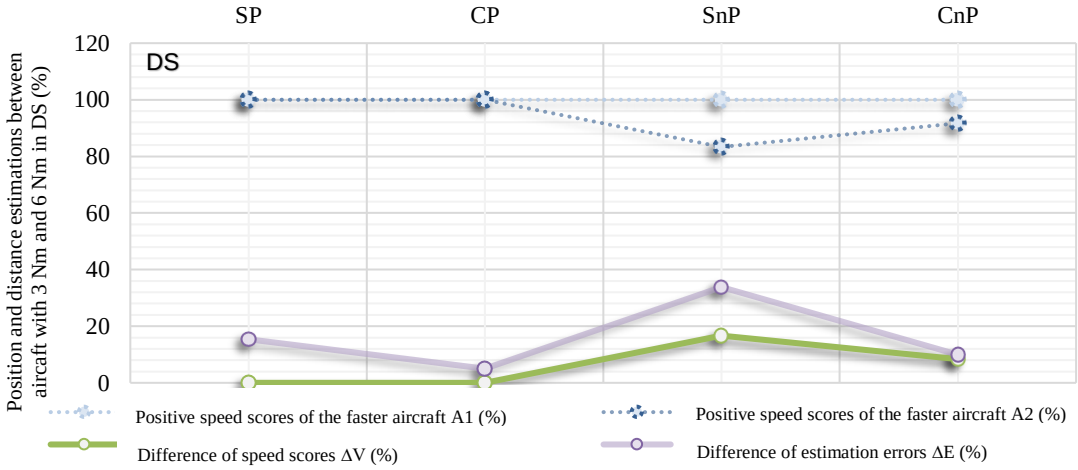


Figure 51: *Position and distance estimation differences of the two aircraft (i.e., slower and closer aircraft A1 in blue, faster and farther aircraft A2 in red) relative to the correct future positions for DS stimuli (i.e., aircraft moving at different speeds) with 3 Nm and 6 Nm minimum separation.*

Overall, we observed that participants judged future aircraft positions with smaller position estimation errors and with a more accurate perception of the relative motion between the two aircraft with the CP animations compared to the other animation design types (i.e., ΔE in Figure 49). Participant position estimation errors increased and their accuracy in the perception of aircraft relative motion decreased the least from CP, in the SP, followed by the CnP, and then by the SnP animations. We also found that the aircraft speed scores of both aircraft were positive in all animations design conditions (i.e., speed scores of A1 and A2). Thus, the participants perceived the slower and closer aircraft (i.e., A1) to be much

slower, and the faster and farther aircraft (i.e., A2) to be much faster than their true motion, irrespectively of the animation design type used to estimate future aircraft movements.

In the *continuous animations depicting path history (CP)*, we found that participants judged the future positions of both aircraft in a similar way. As shown in Figure 49, they estimated the positions of both aircraft with similar accuracies (i.e., the difference of estimation errors between A1 and A2 was the lowest one) and they perceived the speeds of both aircraft in a similar way (i.e., the speed scores of A1 and A2 were both positive and their difference was low) than for the other animation designs. This meant that continuous animations displaying path history allowed participants to accurately estimate future positions of aircraft by consistently maintaining the relative motion between aircraft.

In the *semi-static animations depicting path history (SP)*, we found similar patterns in participant estimations as for the CP animations. However, their estimation errors of both aircraft future positions were greater than those of the CP animations (Figure 49). In addition, participants perceived both aircraft speeds to be slower (less positive) than those of the CP animations. In the DS condition, in which both aircraft moved at different speeds, minimum separation distance estimations were shorter than with the ES condition, as shown in Figure 50. This meant that participants erroneously perceived the two aircraft as moving at similar speeds.

CnP animations revealed a greater difference in participant position estimations between aircraft A1 and aircraft A2 compared to the CP and SP animations. As shown in Figure 49, approximately 90% of the participants perceived the faster and farther aircraft (A2) to be faster than the true speed. In contrast, only 60% of the participants perceived aircraft A1 to be faster than the true speed. In the ES condition, the speed of A1 has been estimated to be even slower than the true speed.

SnP animations exhibited the worst performance due to having the greatest difference in participant speed scores and position estimation errors of both aircraft (Figure 49).

In the *ES condition*, in which both aircraft move at equal speed, participants found aircraft A2 to be faster than its true motion, in all animation design conditions (Figure 50). In contrast, the perception of the A1 speed varied considerably among the four designs. Our analyses also revealed that in the animations depicting path history (in the ES condition), participants perceived more accurately the relative motion between the two

aircraft compared to animations without path history. Moreover, these perceptions were slightly more accurate with continuous animations than with semi-static animations.

In contrast, in the *DS condition*, in which A2 moves faster than A1, participants estimated the speed of A1 to be faster than its true motion, in all animation design types (Figure 51). Similar to the ES condition, participants perceived more accurately the relative motion and future aircraft positions with the continuous animation type, compared with the other designs. With semi-static animations, participants judged the positions of A1 with greater positive estimation errors than the positions of A2. They thus perceived the speed of A1 to be faster than its true motion, and the minimum separation distance between the two aircraft to be shorter than the true distance.

To summarize, the depiction of coordinated moving objects with continuous animations allowed participants to easily create a mental model of the future aircraft movement dynamics. In semi-static animations, participants had more difficulty in conceiving correctly the relative motion between aircraft, especially when they moved at different speeds. In this case, participants erroneously conceived both aircraft to be moving at a similar speed. In addition, with continuous animations, participants conceived future movements with mostly positive speed scores, and thus as aircraft moving faster than their true motion, which might be advantageous in air traffic control for promptly detecting conflicts between aircraft. We also found that the depiction of aircraft path history was important in estimating future aircraft dynamics, especially for semi-static animations, in which participants showed the worst task performance. These results confirmed our expectations according to our first specific hypothesis **WH 1** of this experiment.

In the next section, I present the results of self-reported stress-related factors and EDA analysis to better understand the affective state of the participants during the experiment, and thus to better explain their task performance.

6.4.4 Analysis of User-Related Factors

Short Stress State Test Questionnaire (SSSQ) and Affective State

According to the results of the Short Stress State Questionnaire (SSSQ) (Helton 2004), participants were, on average, quite motivated, slightly distressed, and somewhat worried when solving the experiment tasks. Figure 52 shows the averaged SSSQ scores of the three stress factors (i.e., engagement, distress, and worry) before (i.e., pre-scores) and after (i.e., post-scores) the experiment, as well as participants' mean scores for each stress factor. We found a significant difference among mean stress scores between engagement and distress, and between engagement and worry respectively, ($\chi^2(2) = 18.72$, $p = .000$).

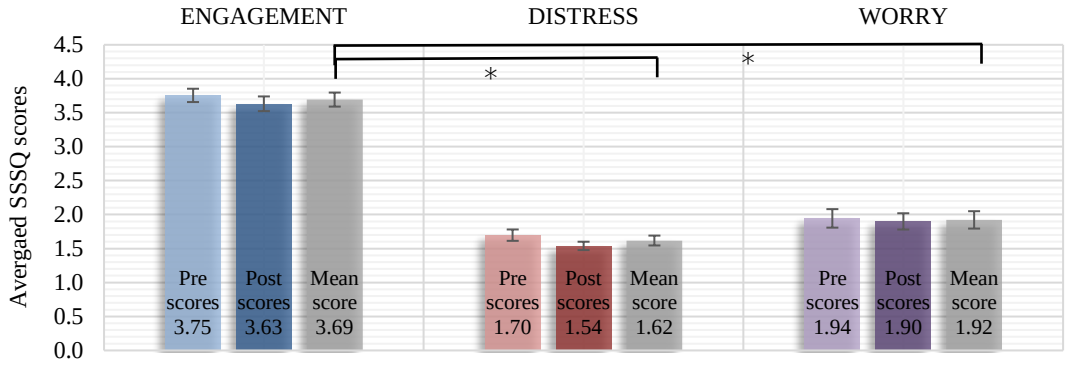


Figure 52: *Averaged SSSQ pre-scores and post-scores for the three stress factors, engagement, distress and worry, as well as their mean values, on a Likert 5-point scale.*

We calculated standardized SSSQ z-change scores for all three stress factors as the difference of the measured stress scores before and after the experiment (i.e., Formula 2). δ corresponds to the standard deviation of the pre-scores (Helton 2004). On average, all the three factors showed negative z-change scores, but these results did not differ significantly, as shown in Figure 53.

$$\Delta S_{z-change} = \left(\frac{S_{post-scores} - S_{pre-scores}}{\delta}\right)$$

Formula 2: *SSSQ z-change scores (Helton 2004).*

Figure 53: *Standardized SSSQ z-change scores across the three stress factors.*

We further compared the SSSQ z-change scores of the three stress factors across two training levels, i.e., low training (< 10 years) and more training (>= 10 years). As illustrated in Figure 54, participants with longer training seemed to be more engaged, more distressed, and less worried than participants with low training, but these results did not differ significantly.

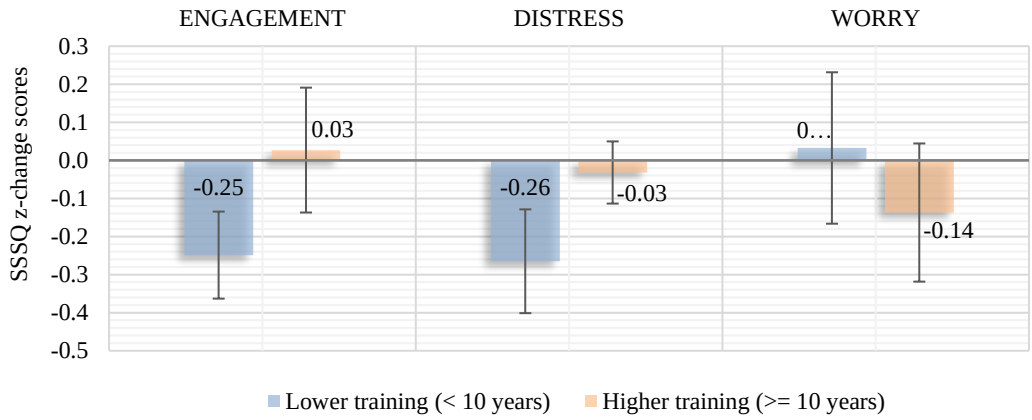


Figure 54: *SSSQ z-change scores across two training levels of the participants.*

The following section provides the results of the analysis of participant electrodermal responses and the comparison between their affective state and task performance.

Electrodermal Activity (EDA)

We calculated the electrodermal activity (EDA) of participants as the mean integral values of their normalized phasic skin conductance responses (i.e., the *area under the curve*, AUC) per stimulus and per participant (Boucsein 1992; Figner & Murphy 2010). As for performance accuracy, we normalized the AUC scores for each participant according to Formula 3:

$$AUC_{norm,perc} = \frac{AUC * 100}{AUC_{max} - AUC_{min}}$$

Formula 3: *Normalized AUC scores (Lykken 1972).*

This normalization method is recommended for reducing the effect of age, gender, and other potentially influential individual differences, and thus enabling the comparison of AUC values across participants (Lykken 1972). The EDA data only included nine participants, because the data of three participants was not correctly recorded.

We compared normalized mean AUC values (in %) of the four animation design types. Participants' AUC scores did not differ significantly across animation design types, as shown in Figure 55. Normalized mean peak counts (in %) showed similar patterns as the computed AUC values across animations design types. The higher the intensity of the electrodermal activity is (i.e., high AUC values), the more physiological reactions were found (i.e., mean peak counts). Peak counts did not show any statistically significant differences across animation design types.

Further, we analysed AUC value changes according to the stimuli order presented to each participant. We found that in 66.7% of the cases, participants' AUC scores decreased over time. This result identifies a correspondence with the SSSQ z-change scores, in which the sum of the three stress factors showed a decreasing trend in 6 of 9 (i.e., 66.7%) participants. The chi-squared statistic of independence of the two categorical variables on nominally scaled data (i.e., -1 for decreasing and +1 for increasing AUC and SSSQ scores) showed that the participants' electrodermal activity is meaningfully interconnected with their affective state ($\chi(1) = 5.14$, $p<.023$). This meant that when participants' self-reported stress state decreased, their correspondent electrodermal activity decreased, as well. This relationship is shown in more detail in the next section.

Finally, we also compared averaged EDA responses across training levels, but we did not find any significant differences.

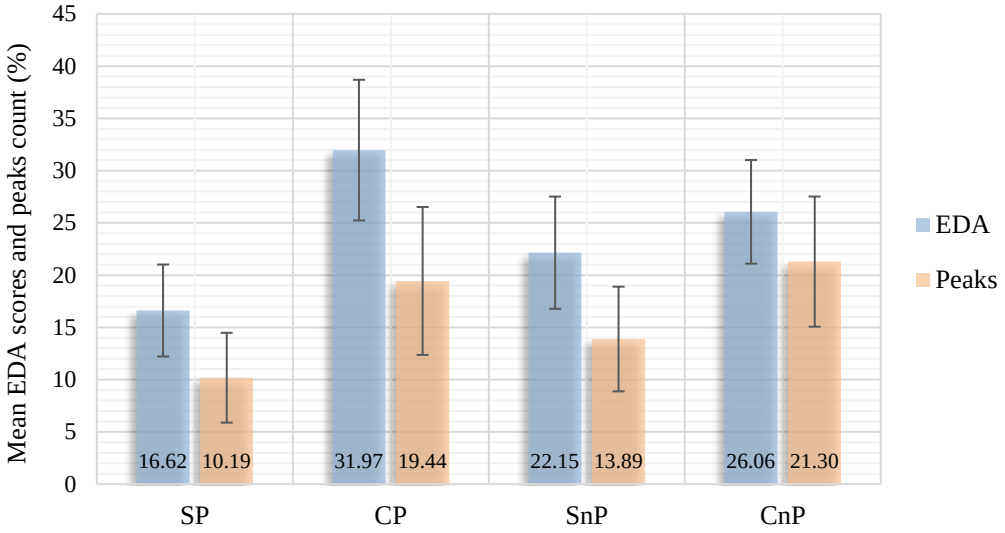


Figure 55: *Participants' AUC mean values (%) and mean peaks count (%) across the four animation designs.*

Comparison between Task Performance, EDA, and Affective State

How task performance might be related to electrodermal activity and self-reported stress states (i.e., engagement and distress) is discussed below. First, we correlated participant physiological responses with their task performance, and second, we compared this with participant scores on the self-reported SSSQ questionnaire.

Considering averaged EDA values and task performance across the four animation design types, we found similar patterns between the two dependent variables (Figure 56). More accurate responses were associated with higher physiological arousals. Interestingly for semi-static animations with path history (SP), while participants showed good performances, their EDA was lowest compared to the other three animation design types.

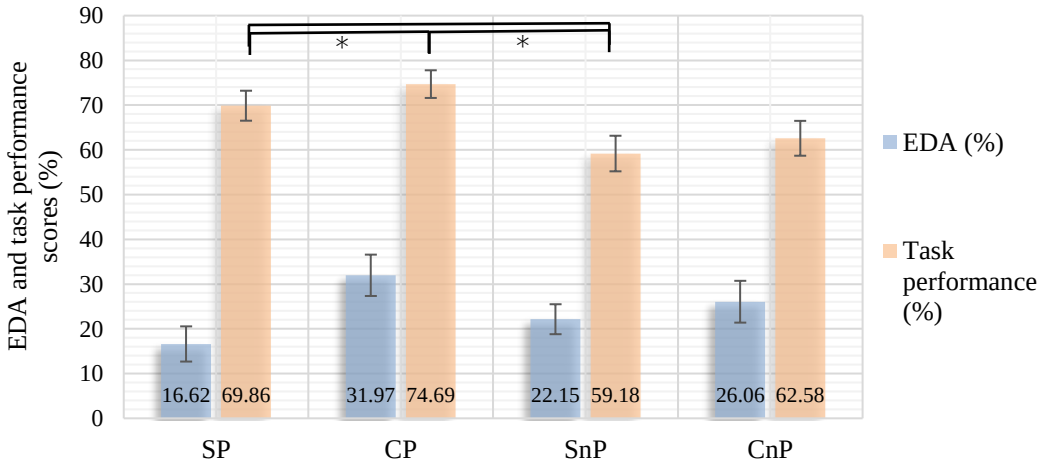


Figure 56: *EDA compared with task performance across animation design types.*

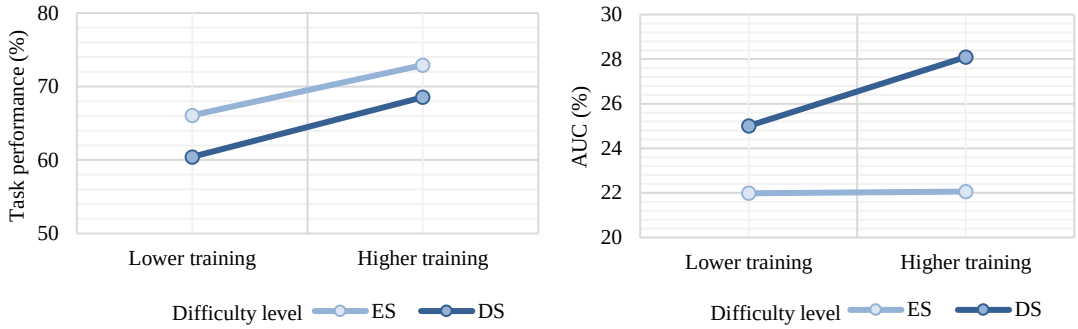


Figure 57: *Task performance (left) and EDA (right) across difficulty (i.e., ED vs DS) and training.*

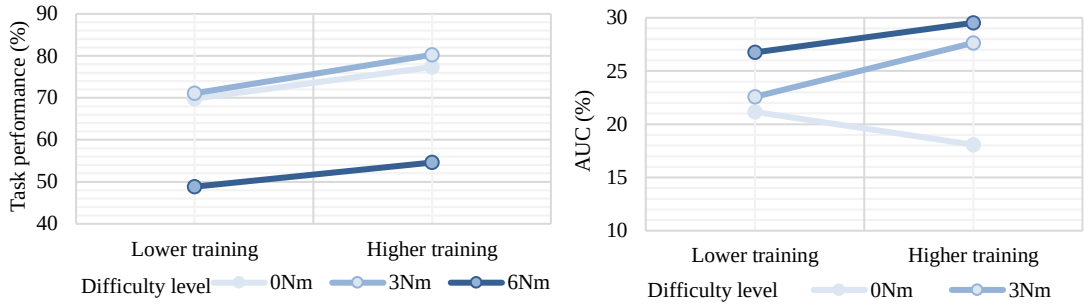


Figure 58: *Task performance (left) and EDA (right) across difficulty (i.e., 0-3-6 Nm) and training.*

As hypothesized (cf. **WH 3.1**), we observed an increase of the task performance and a decrease of the electrodermal activity by decreasing the task difficulty (i.e., ES vs DS, and 0 Nm vs 3 Nm and vs 6 Nm) (Figures 57 and 58). Not surprisingly, by increasing the training level (i.e., low vs more training) of the participants, their task performance increased and electrodermal activity increased as well (Figures 57 and 58).

Further, we summed the SSSQ z-changed scores of two stress factors, i.e., engagement and distress, to correlate participant affective states with their response accuracies. We calculated the bipolar valence of affective state (i.e., negative or positive state) by subtracting distress scores from engagement scores. According to Boyle et al. (2014) and Matthews et al. (2013), we assume engagement is a positive affective state, whereas distress is negative (cf. Chapter 2, *Related Work*). The engagement, distress, and their associated bipolar valence scores for all 9 participants are presented in ascending order in Figure 59.

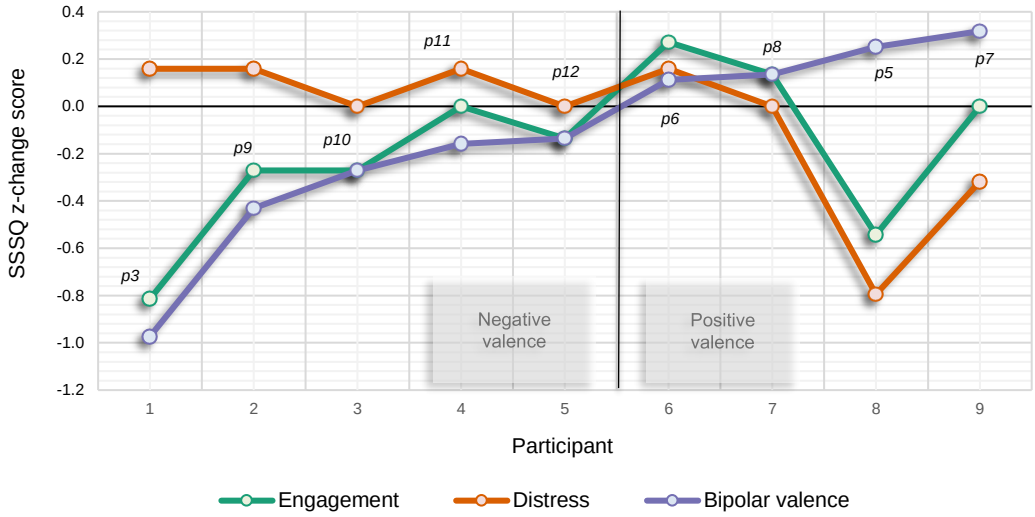


Figure 59: *Standardized z-change scores of the SSSQ responses across the three stress factors and the corresponding bipolar valences.*

According to Russell's *circumplex model of affect* (1980) and Larsen and Diener's *compromise circumplex* (1992) (cf. Chapter 2, *Related Work*), task performance, EDA, and bipolar valence should be correlated to one another. We qualitatively illustrate this relationship between bipolar valence and EDA scores of the participants in Figure 60, in which we also add their corresponding task performance ranking (i.e., 1 for the highest response accuracy to 9, the lowest response accuracy) and training level (i.e., HT for more training and LT for low training level). As presented in Chapter 2, *Related Work*, high arousal values and positive emotion valences may be related to higher performance scores, whereas lower performance scores may be related to low arousals and negative valences. In addition, longer-trained participants showed higher arousals, higher positive valence (i.e., engagement and calm), and higher task performance levels. Conversely, low trained participants exhibited lower arousals, higher negative valence (i.e., less engagement and boredom), and performed less accurately.

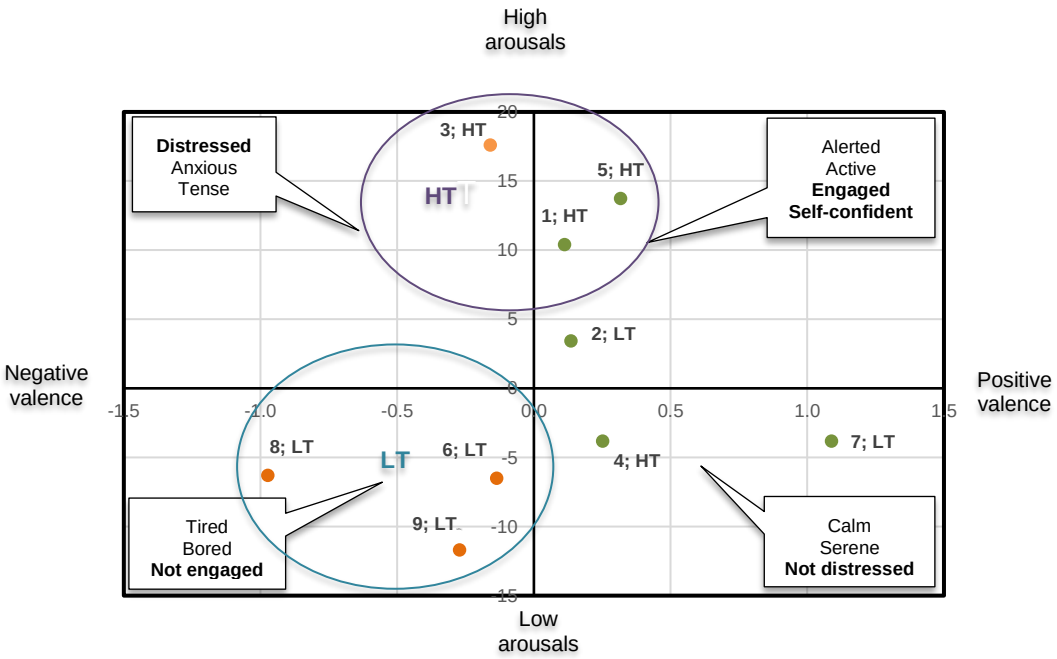


Figure 60: *Plot of each participant's arousal intensity versus emotion valence. Each participant's task performance ranking (i.e., from 1 to 9) and training level (i.e., HT and LT) are also noted. Green points correspond to positive affective states (e.g., engagement and calm) and orange points to negative affective states (e.g., boredom and distress).*

6.5 Key Findings and Discussion of Experiment II

In our second user study with air traffic controllers, we investigated how *task difficulty* (i.e., different minimum separation distances between moving aircraft and relative speeds), *animation design* (i.e., semi-static vs. continuous animations, and displays depicting path history vs. displays without path history), and *user-related factors* (i.e., training level and stress) might influence predictions of future aircraft movements. The goal was to better understand how controllers perceive the relative motion of aircraft, and what decision-making factors influenced ATC task performance (i.e., estimation of minimum separation distance between two converging aircraft). The results showed that controller task performances were significantly affected by the animation design type and by the task difficulty, in this case, the motion complexity of the depicted aircraft. The results were also influenced by user-related factors. The following three sections explain in more detail the implications of the studied factors for the prediction of future aircraft movements, a critical ATC task.

6.5.1 Context of Use and Task

We tested the effect of two task difficulty levels on the prediction of future aircraft movements with animations. The first level involved the estimation of future aircraft movement patterns at three different minimum separation distances. The second level concerns the estimation of future aircraft movement patterns while moving at similar and at different relative speeds. The task has been designed according to the third SA task (Endsley 1995; Grier 2015), i.e., how effectively viewers project near-future movement patterns, and with realistic ATC displays. Our hypothesis, mentioned in Section 6.2 , *Specific Research Questions and Working Hypotheses*, was as follows:

> ***WH 1:*** *Task performance decreases with an increase of the task complexity, i.e., by presenting animations displaying aircrafts with different relative speeds and with increasing minimum separation distances.*

We found that differences in the relative motion configurations of two converging aircraft, i.e., by changing their relative speeds and minimum separation distances, influenced task performance. As hypothesized (cf. **WH 1**), we find that controllers, on

average, performed the prediction task significantly less accurately with displays showing aircraft at different speeds compared to displays showing aircraft at same speeds. Similarly, controllers' average accuracy decreased by increasing minimum separation distances between two converging aircraft. This finding confirms previous studies in the ATC context with semi-static visualizations (Rantanen & Nunes 2005; Boag et al. 2006). As mentioned in the first section of this chapter (cf. Section 6.1*, Motivation and Research Goal*), we observed that controller performance decreased by increasing the relative speed difference and/or the distance of minimum separation between aircraft with semi-static animations. However, on average, participants estimated in similar way minimum separation distances of 0 and 3 Nm. Their estimations differed significantly between distances of 0 and 6 Nm, and between 3 and 6 Nm. Participant estimations of future aircraft movement patterns at 0 and 3 Nm were around 75% of correct responses, and decreased to 51% with a minimum separation distance of 6 Nm.

This result might be explained by training. The NSEEV model (Steelman et al. 2011) highlights that training is important for the correct interpretation of the criticality and urgency of specific air traffic conflicts in supervisory visualization tasks. Controllers are especially trained in effectively and efficiently identifying those aircraft that move too close to one another, i.e., below a minimum separation distance of 3 Nm (minimum safe distance on the horizontal plane). All other distances above this minimum safe distance, such as 6 Nm, do not imply an air traffic conflict. This might explain the better performance of controllers in predicting aircraft movement patterns at critical air traffic situations compared to situations without conflicts.

The effect of motion complexity on the prediction of future aircraft movements differed significantly across the four animation design types. With continuous animations depicting path history, the effect was less strong than with the other three animation types, i.e, with semi-static animations with and without path history, as well as with continuous animations without path history. This suggests that controllers even with the new continuous animations could effectively perceive the relative motion, and thus effectively predict distances of minimum separation, between two converging aircraft. This finding provides empirical evidence and confirms contentions in previous animation research in cartography and GIScience, that continuous animations are well-suited for effectively displaying and perceiving relative motion between moving objects (not yet empirically tested), and thus for effectively predicting near-future movement patterns (Griffin et al.

2006; Schlienger et al. 2007). This means that continuous animations might be appropriate in cartography and GIS (as well as in other domains such as sport analysis or transportation system) to depict movement data and relative motion, especially for experts that need to quickly recognize specific motion patterns of one moving object in relation with another one. Important applications in cartography may be dynamic geographical data, spatio-temporal data such as hurricanes, migration of people etc. Applications in sport may include analysis of a soccer match; in transportation systems, analysis of anomalous situations, such as when a train has been delayed.

Next, the effect of animation design on the prediction of future aircraft movement patterns is discussed.

6.5.2 Animation Design

We tested the effect of animation design on the prediction of future aircraft movements by manipulating the dynamic visual variable 'rate of change' of the animated displays (i.e., semi-static vs. continuous animations), and the depiction of aircraft path history (i.e., displays depicting path history vs. displays without path history). Our hypothesis concerning animation design, mentioned in Section 6.2, *Specific Research Questions and Working Hypotheses*, was as follows:

> ***WH 2:*** *Continuous animations and path history improve task performance by helping controllers to easily perceive relative motion and thus to easily predict future aircraft patterns (compared with semi-static animations and displays without path history).*

The effective perception of *relative motion* between aircraft is important for air traffic controllers to perform one of their most important tasks, i.e., to accurately estimate the minimum separation distance between two converging aircraft. We observed that controllers apprehended the relative motion between aircraft more effectively with the novel continuous animations, compared to standard semi-static animations, irrespective of the depiction of aircraft's path history. The continuous animation might thus allow controllers to better track the movement dynamics of one aircraft in relation to others on the display, and, consequently, be able to better forecast future movements and locations of two aircraft. We thus can, as hypothesised (cf. **WH 2**), demonstrate that continuous animations may be effective and beneficial for more effectively apprehending relative

motion patterns, as well to better predict near-future movement patterns, in this case, of two coordinated moving aircraft. This result is in line with the *Common Fate Principle* (Koffka 1935), and the *Event Perception Theory* (Shipley & Zacks 2008; Shipley et al. 2013), which support the idea that users do not treat spatio-temporal phenomena or events in isolation, but in relationship with their context.

Moreover, we found that the *depiction of aircraft path history*, as a standard representation of aircraft past positions in ODS displays (never empirically tested to date), did indeed support controllers in effectively tracking aircraft dynamics and, thus, in better predicting future aircraft movements. Previous studies on animations (Robertson et al. 2008; Slingsby et al. 2010) emphasize the effectiveness and efficacy of *traces* to analyse spatio-temporal trends compared to animations without any path information. However, Robertson et al. (2008) empirically found that traces on animations are more suited to visualizing small amounts of data. With larger datasets, traces tend to be confusing because of visual clutter. As ATC screens only show at most 25 moving aircraft, the added graphic marks do not clutter the screen unnecessarily.

In conclusion, the first hypothesis of our second experiment (cf. **WH 2**) could not be rejected. Continuous animations, and especially their combination with additional visual features like path history, allowed users to effectively create a mental model of the current moving object dynamics and, consequently, to better predict future movement patterns. However, task performance might be influenced by specific user characteristics as well. The next section will cover this in more detail.

6.5.3 User Behaviour

According to the *Yerkes-Dodson* (Yerkes & Dodson 1908) and *Csikszentmihalyi* (1990) theories (cf. Chapter 2, *Related Work*), task performance might be influenced by different internal and external factors, such as user stress, motivation, positive feelings, training, and the perceived task difficulty. Concerning user-related factors, we derived the following hypotheses:

> ***WH 3.1:*** *Task performance increases with a decrease of stress factors (i.e., more distress, more worry and less engagement) and of task difficulty, and respectively with an increase of training and EDA values.*
>
> ***WH 3.2:*** *Familiarity with semi-static displays, and with displays showing path history, influence positively controller EDA values, design preferences and ease of use ratings. We (and they) will expect to perform better with semi-static displays depicting path history than with the other animation design types.*

As hypothesized (cf. **WH 3.1**), we found that controller training level, their stress states, and perceived task difficulty might influence their task performance.

First, we found that controllers with more training (≥ 10 years of training) performed the tasks better compared to controllers with less training (< 10 years of training). As expected (cf. **WH 3.1**), this difference in their task performance was also reflected in their emotional states during the experiment as well. Controllers with more training not only performed the task better, but they also showed higher EDA values, compared to controllers with less training. This is in line with the *Yerkes-Dodson law* (1908), which states that task performance has a curvilinear (or also called inverted-U) relationship to electrodermal activity. Task performance and electrodermal activity increase to a critical level, when stress and task difficulty are too high. After this critical point, both performance and electrodermal activity decrease.

Furthermore, participants also showed differences in the tested stress-related factors (i.e., engagement, distress, and worry). Controllers with more training seemed to be more engaged, more distressed, and less worried than controllers with less training while solving the experiment task. This increase of positive affective states (engagement), and the decrease of negative affective states (distress and worry) respectively, might have positively

affected their task performance. This is in line with previous research concerning the psychology of stress. For example, Boyle et al. (2015) and Matthews et al. (2013) empirically demonstrated that task performance may be correlated with engagement and distress. In turn, Matthews et al. (2013) also suggest that engagement might be associated with attentiveness, joviality, and low fatigue, whereas distress may be linked to low serenity and anxiety. In Russell's *circumplex model of affect* (1980), the type of affective state felt at a certain moment can be derived by associating the intensity of psychological arousals (i.e, high arousals vs. low arousals) with the valence of the affective state (i.e., positive vs. negative affective state). For example, people showing high arousal and positive valence might mean that they are engaged and/or calm, whereas people with low arousal and negative valence might be a sign of fatigue and/or boredom.

Finally, controller familiarity with standard semi-static animations, and with displays showing path history, might also have an influence on their EDA. However, this was not as we expected (cf. **WH 3.2**); we found that controllers' EDA with semi-static animations were lowest compared to the other three tested animation types, whereas with novel continuous animations the electrodermal intensity was higher than with semi-static animations. According to previous studies in neuroscience (van den Bosch et al. 2013), this is unusual. EDA should positively correlate with respect to familiarity and self-reported pleasure using the displays under study. This unexpected result might be explained by the self-reported preference ratings of participants. Most of the controllers reported that they preferred to solve the task with continuous animations compared with semi-static animations. They believed that continuous animations allowed them to easily predict future aircraft movements, because of the uninterrupted motion of the aircraft dynamics. Consequently, in contrast to the conclusions of Lee and Klippel (2005), this result demonstrated again (cf. Section 6.4, *Results*) that the performance of experts might not be negatively influenced by novel or unfamiliar display design. Our findings suggest that controllers' heuristics gained by training with a specific graphic display did not impede them from processing information efficiently and effectively with new kinds of visualizations. In our case, experts not only preferred working with continuous animations, but they also performed better with them. Perhaps continuous animations that congruently depict the smooth movement of aircraft might allow them to more easily create a mental representation of actual aircraft movement dynamics. In other words, when high quality information is available to ATC experts, they can take advantage of it.

To summarize, according to our results and our hypothesis (cf. **WH 3.1**), we can derive that users with more training showed higher physiological arousals, stronger positive emotions (engagement or calmness), and more accurate task performance. Conversely, users with less training showed lower physiological arousal, stronger negative emotions (distress or boredom), and less accurate task performance. This finding may have a direct impact in the ATC context. Measuring physiological arousals of controllers during their work might be useful, for example, to more easily recognize stress-related symptoms, such as fatigue or anxiety, and, consequently, to react in case of stressed situations. For example, in the case of ATC, Costa (2001) proposes to reduce working hours or increase pauses according to working load and self-perceived stress. This author also suggests to ameliorate the ergonomic design of interfaces with respect to self-perceived or external (e.g., noise) stress factors. Developing appropriate responsive display designs in real-time according to the affective states of controllers may also help in reducing their stress levels and in increasing their task performance. In addition, we also found that familiarity with standard displays might not impede controllers in solving ATC tasks effectively and efficiently with novel types of animations. In our case, controllers preferred and performed the task better with novel continuous animations than with standard semi-static animations. This finding might encourage system designers in ATC, and more in general in cartography, GIScience and in other application fields where movement data is displayed, in developing appropriate visual animated displays according to user-related factors (such as their preferences and their affective states) to improve task performance.

In the next chapter, I summarize and compare the previously mentioned key findings of the two experiments to derive general design recommendations for animations depicting movement data (cf. Section 5.5, *Key Findings and Discussion of Experiment I* and Section 6.5, *Key Findings and Discussion of Experiment II*).

CHAPTER SEVEN

GENERAL DISCUSSION

In this thesis, I have presented two empirical studies with animated displays. Both experiments were designed and realized in the context of air traffic control (ATC) with simulated ATC radar displays and realistic Situational Awareness (SA) tasks, i.e., (1) in the *apprehension of aircraft movement changes*, and (2) in the *prediction of future aircraft movement patterns*. We investigated the influence of the three main visual analytics (VA) dimensions, i.e., the *use context and respective task types* (i.e., SA tasks with different task difficulty levels), the *animation design* (i.e., semi-static vs continuous animations; displays with/without path history), and *user-related factors* (i.e., expertise, affective states, spatial abilities, etc.), on the effectiveness and efficiency of decision-making with animations. Our results suggest that these three VA dimensions indeed influence users' decision-making processes with animated displays. This section first summarizes the main outcomes of both experiments to answer our general research question (cf. Section 1.5, *General Research Questions*):

How should animations of real-time movement data be designed considering the task and/or use contexts, and user characteristics?

The main outcomes of both experiments are discussed according to the three research sub-questions (i.e., RQ 1, RQ 2, and RQ3) mentioned in Section 1.5, *General Research Questions*. Subsequently, critical arguments and limitations of this research are discussed.

Finally, a list of general recommendation to improve the animation design of movement data to achieve more effective and efficient visuospatial decision-making with animations.

7.1 Context of Use and Task

As presented in Section 1.5, *General Research Questions*, our first sub-research question of this thesis is the following:

> ***RQ 1:*** *How does the use and task contexts influence user visuospatial decision-making with animations depicting movement data?*

We identified the context of air traffic control (ATC) as an appropriate and real-world use case for effectively testing animations. In our two user studies, we investigated (1) how users apprehend and detect aircraft movement changes, and (2) how they predict future aircraft movement patterns and the relative motion between two converging aircraft with animations.

We found that two kinds of task difficulty influence user task performance: (1) the number of the depicted moving aircraft, and (2) differences in their relative motion. The number of the depicted moving entities influences mainly novices, with experts influenced to a lesser degree. In line with Ware (2013), novices more easily noticed movement changes if only a few elements (maximum of 4) were visualized. In addition, similar to Lautenschütz (2011) we found in both experiments that contextual information (i.e, the movement patterns of the other moving aircraft depicted on the display scene) was relevant to identify basic movement changes, such as speed changes, and to predict future movement patterns. However, this might depend on the data complexity. Animated displays depicting few data items (4 elements) and contextual information seemed to be irrelevant. If animated displays depicted a greater amount of data (e.g., 8 moving objects), contextual information became more important. When the number of the depicted objects is increased, users may compare moving objects to one another more often as a strategy to overcome the increasing cognitive demand.

However, there was a substantial difference in the strategies between experts and novices. On average, experts are trained to monitor simultaneously 8–12 aircraft in their everyday job at operational level. Their strategies, different compared with those of novices,

are reflected in our eye movement analysis, as well. We found that they fixated upon all the depicted objects more homogeneously than novices, regardless of the distance and speed difference between aircraft. However, with continuous animations depicting 8 elements, experts often compared objects moving at similar speeds, whereas novices focused their attention more on perceptually salient data (fastest objects or objects closer to each other). The perceptual strategy of the experts in continuous animations confirmed previous findings (Cavanagh & Alvarez 2005).

Finally, similarly to previous studies in ATC (Rantanen & Nunes 2005; Boag et al. 2006), we found that the accurate perception of relative motion between aircraft to predict future aircraft movements depended on the relative distance and relative speed difference between moving objects.

7.2 Animation Design

As presented in Section 1.5, *General Research Questions*, our second sub-research question of this thesis is the following:

> ***RQ 2:*** *Which animation design characteristics might be particularly useful to depict movement data for efficient and effective visuospatial decision-making with animations?*

To answer the above-mentioned research question, in our two user studies we first investigated how aircraft movement dynamics is perceived with two kinds of animated design types, i.e., semi-static and continuous animations. We thus analysed the influence of the dynamic variable *rate of change* (i.e., also smoothness of the transitions between display scenes: abrupt vs smooth transitions) on users' decision-making capabilities with animations (Battersby & Goldsberry 2010). Successively, we also examined the effect of the depiction of path history in inference processes with animated displays. In both experiments, results showed that user task performance with animations might be significantly influenced by animation design in a number of ways, depending on the task type and on user characteristics.

The dynamic variable 'rate of change' affects users task performance differently depending on the task. More specifically, for *apprehension tasks*, we found that continuous animations (i.e., smooth transitions) are not superior to semi-static animations (i.e., abrupt

transitions) in the effective and efficient detection of aircraft accelerations. Perceptually salient objects (i.e., the fastest moving aircraft), which move continuously on animated displays, are perceived in a more salient way than those moving abruptly with a screen refresh rate of 4 Hz. This hinders users, especially novices, from effectively detecting thematically relevant objects (i.e., aircraft accelerations). This finding confirms previous studies on animations (Boucheix & Lowe 2010; Fabrikant 2005), and is in line with the *Event Perception Theory* (Shipley & Zacks 2008) and the *Gestalt Theory of common fate* (Koffka 1935). In addition, this finding seems to be influenced by other factors such as user characteristics (i.e., expertise, familiarity with the displays and task, affective state, and spatial abilities of the users). For example, experts performed well with both continuous and semi-static animations, while novices performed significantly better with semi-static animations than with continuous ones.

A possible explanation is that we only tested viewer's ability in detecting speed changes (i.e., accelerations) and not direction changes of the aircraft motion under study. The particularity of aircraft accelerations visualized on standard radar displays is that they occur very slowly. How effectively accelerations are perceived by the viewers depends on the animation design type. With continuous animations, movement changes are visualized by means of gradual and smooth transitions between scenes, while with semi-static animations, they are visualized by means of abruptly refreshing scenes. We believe that, with continuous animations, slow speed changes, such as real aircraft accelerations, might be too smooth to be effectively perceived. Similarly, Fabrikant (2005) warned that transitions between dynamic scenes (i.e., tweening) might be useful to mitigate the change blindness effect, but excessively smooth transitions might make it more difficult to effectively detect aircraft speed changes. The question that arises is if smooth transitions between scenes are more suited to visualize other kinds of movement patterns, such as aircraft direction changes. As highlighted by Mateeff et al. (2000), humans are more visually sensitive to direction changes than to speed changes (e.g., accelerations). Consequently, aircraft direction changes that were shown on an ATC animated display might thus have been more easily noticeable compared to aircraft accelerations. In self-reported questionnaires and interviews, some experts reported that direction changes were relevant movement changes to be detected in their job and that continuous animations might be more effective than semi-static ones to be more promptly and accurately perceived. With continuous animations, they might more easily track, and without interruptions, how the

movements' changes (e.g., the change rate and magnitude of the direction change) develop over time. However, this has not been tested empirically to date.

In contrast, for tasks involving the *prediction of future aircraft movements* (i.e., future aircraft positions and distances) of a pair of converging aircraft, participants performed more accurately with continuous animations compared to semi-static animations. As air traffic controllers expected, the continuous representation of the movement dynamics allows them to accurately perceive the relative motions between moving objects, and thus to effectively and efficiently predict future aircraft movements. Moreover, participants reported in post-test questionnaires that they preferred the continuous animations over the semi-static ones for solving the required tasks. The uninterrupted motion of aircraft allowed them to access more detailed information about the dynamics and relative speeds of aircraft compared to semi-static animations. With continuous animations, they did not have to wait 4 seconds until acquiring information about the next aircraft movement information. The brain data analysis of the first experiment showed as well that participants were more positively motivated (higher positive approach-related motivation) to perform the task with continuous animations than with semi-static animations. This confirms previous studies that users find animations to be more enjoyable, easier, and more exciting to use compared to other kinds of visual displays (Kriglstein et al. 2012).

We also analysed the cartographic visual variable *path history* (or trace) on the prediction of future aircraft movements. Path history is used to depict aircraft past positions on standard French ATC displays (i.e., semi-static animations), but no empirical evidence concerning its efficacy and effectiveness in predicting future aircraft movements exists in this domain to date. In past animation studies (Robertson et al. 2008), trace visualizations have been mostly empirically tested for data presentations, but not for data exploration and analysis. We thus investigated the importance of path history for air traffic controllers to keep track of aircraft movements in animations. We found that the visual variable *path history* might indeed be an appropriate variable to depict small quantities of movement data (<25 entities) used for tracking movement objects.

To summarize, continuous animations are adequate for users in performing the third SA task, i.e., predicting near-future aircraft dynamics. Because of the smooth transition scenes permitted by continuous animations, aircraft movement patterns are perceived in a more realistic and natural way. This allows users to better perceive the relative motion

between moving features. Conversely, no superiority of continuous animation relative to semi-static animations has been found in the detection of generic movement patterns (i.e., speed changes) (Dodge et al. 2008). Especially for novices, accelerations were probably too slow and too smooth to be effectively identified with continuous animations. Other factors might influence task performance, and some of them might help us to explain why some participants performed better than others. In the next section, relevant user characteristics, e.g., expertise and familiarity of the user, their motivation, their affective state, and their spatial abilities, that might influence decision-making processes with animations are discussed.

7.3 User Behaviour

As presented in Section 1.5, *General Research Questions*, our third general research question of this thesis was the following:

> ***RQ 3:*** *How do viewers' perceptual, cognitive, and affective states influence the effectiveness and efficiency of visuospatial decision-making with animations depicting movement data?*

Previous human-computer and geo-visualization studies suggest that user prior knowledge and visual spatial skills influence the effectiveness and efficiency of information processing with animations (Fabrikant 2005; Kriz & Hegarty 2007). In our first experiment, we also found that participant task performance (i.e., response accuracy in detecting aircraft movement changes) might be affected by their familiarity, training, and spatial abilities. In addition, other user-related factors, such as the affective state of the participants, seem to influence decision-making processes with animations.

First, we find that *expertise and familiarity* with the tested displays and tasks significantly influence how users perceive and make decisions with animations. As discussed in the previous section, domain novices might be more influenced by bottom-up processes (i.e., the perceptual salience of the depicted features) than domain experts. For *experts, in contrast to novices, top-down processes* (i.e., their prior knowledge) seem to be more relevant than bottom-up processes in visual information processing with animations. In the first experiment, experts detected aircraft movement changes well with both novel continuous

and familiar semi-static animations (no significant difference in their response accuracies was found). In our second experiment, we found that experts not only performed better with continuous animations, but they also preferred to solve the task using these compared to familiar animated displays. This means that familiarity with semi-static animations might not significantly influence the task performance of experts with unfamiliar animated display. This result is in contrast to the contention of Lee and Klippel (2005), who assume that if air traffic controllers use unfamiliar design types (continuous animations) to monitor aircraft movements, even if well-suited to display movement patterns and changes according to the Congruence Principle (Tversky et al. 2002) and Gestalt Theory (Koffka 1935), they might mismatch with their heuristics resulting in worse performance.

The above-mentioned findings may be explained by participants' *spatial skills* as well. We found that superior *spatial skills* might help users to perform better in visuospatial tasks with animations than low-spatial users. As highlighted by Newcombe and Frick (2010), training in specific visuospatial tasks with animations, such as well-trained air traffic controllers monitoring aircraft on dynamic radar screens, might improve users' spatial skills, as the significant difference in performance seen in the Hidden Patterns Test between experts and novices shows. Well-trained users (experts) or novices with high spatial skills are then facilitated in the information extraction from not only well-known displays, but also novel and unfamiliar visualizations, by transferring their prior knowledge and learned mental models with higher neuronal efficiency and lower cognitive load (Grabner et al. 2003).

Moreover, according to our EEG analysis, user *motivation* in solving the experimental task might affect their task performance, as well. Surprisingly, we find that novices were more motivated than experts in solving the task. Past studies in maintaining vigilance on ATC monitor tasks argue that this might be due to the monotony of the tasks for experts, as these tasks might be too similar to their everyday job (Matthews et al. 2013). In contrast the novelty of the task for novices might arouse their motivation (Barto et al. 2013). From the post-test questionnaires, some novices reported that the task was similar to a videogame. As reported by Przybylski et al. (2010), playing video games leads to positive affect resulting in increasing motivation of goal-directed behaviour and in the activation of positive emotions. As expected, the motivation levels of both experts and novices were higher when exposed to continuous animations rather than to semi-static ones. As

Kriglstein et al. (2012) emphasize, continuous animations are often judged as more enjoyable than other visualization types, such as static displays.

In both experiments, the *electrodermal activity* (EDA) of well-trained experts showed higher electrodermal activity compared with novices and less-trained experts. Well-trained experts were also more engaged, less stressed, and less worried that novices, and less-trained experts. These findings are in line with Russell's *circumplex model of affect* (1980) and Csikszentmihalyi's *flow theory* (1990), both of which state that participant positive affective states (higher engagement, respectively lower distress and worry) positively affect task performance. We also find that participant arousal is positively correlated with task performance (i.e., response accuracy). In addition, users may have a sense of how well they are doing, and this may influence their affective state (i.e., engagement, distress, and worry), and perhaps their cognitive state (e.g., motivation) as well. There might be an interaction with their affective/cognitive states and their expertise. Controllers with more training (or more generally, 'experts') may have a better sense of how well they are doing, and have a different response to doing well, respectively poorly, relative to controllers with less training (or more in general novices). This might influence positively, respectively negatively, to users' task performance.

Participants were more aroused by continuous animations than with semi-static animations. In the first experiment, this might have been due to the higher cognitive load induced by the continuous processing of information. This was also confirmed by anecdotal evidence, as participants reported that, with this kind of animation design, the task was more difficult to solve. However, in the second experiment, experts more easily predicted future movement patterns with continuous animations compared to semi-static animations. In this case, higher arousals might be explained by the novelty of the animation display type and by engagement induced by the greater easiness to solve the task, as suggested by Weierich et al. (2010).

7.4 Methodological Contributions

Our empirical method combining eye movement, EEG, and EDA data with questionnaire and usability metrics to analyse the effectiveness and efficacy of animated displays in decision-making processes is novel in cartography and GIScience. In our empirical studies, standard *quantitative usability measures* (response accuracy and response time), *combined with qualitative metrics* (e.g., interviews and open questions on post-test questionnaires), were effective to analyse and interpret users' task performance, to better understand the influence of users' characteristics on their performance, and to compare different animation designs. In addition, measuring *motivation and cognitive load* through EEG analysis facilitated the explanation of good or bad task performances of the participants. Moreover, considering and measuring viewers' affective states (i.e., engagement, distress and worry) and motivation during an experiment might be an additional source of information to make animations more engaging.

The triangulation of eye movement data, EDA data, EEG data, and SSSQ questionnaire was a strong methodology to evaluate results at higher confidence and reduced uncertainty regarding user decision-making with animations, especially to easily identify the type of cognitive and emotional processes involved. With animations, the strategies adopted by users to identify relevant information often changed over time. This cross-validation approach, which combines different data sources, not only offers an opportunity for the future to use variation in the user-state data (EEG and EDA) to predict performance on a specific trial, but it also allows the analysis of the way in which users process visuospatial information on an animated scene over a certain time period, and how their strategies change over time before making a specific decision. However, this empirical methodology is limited by different factors that I detail in the next section.

7.5 Limitations of the Empirical Methods

Despite the methodological contributions that this thesis has brought about, I should also point out some important critical points and limitations related to the described empirical methodology, that might affect our research, as follow.

The two experiments have been conducted only in a specific *user context* (i.e., in ATC). This makes the generalization of the empirical findings to other similar domains, such as in sport or transportation systems research, difficult. It would be interesting to run the

experiments in other application domains to be able to generalize the results, to improve reliability, and to create sound design guidelines valid for all movement data types and user contexts.

Controlled experiments are suited to isolate the influence of the tested independent variables. The disadvantage of this method is that tasks and information depicted on animations should be sufficiently simple to measure the effect of only the tested independent variables (and thus to control and avoid the influence of undesired factors). In addition, only a few variables can be analysed at once. Consequently, controlled experiments imply a simplification of user contexts and real-life tasks (which are often more complex). For example, in our experiments, we reduced the amount of visual information depicted usually on real ATC animated displays (e.g., we removed text labels, number of aircraft, etc.) and the complexity of tasks that air traffic controllers have to perform simultaneously in their everyday job. It is thus not clear yet, if our findings might be similar in a real context.

Standard methods to analyse *eye-tracking data* of static displays are less suitable for animations. In animation studies, the temporal aspect of the visuospatial decision-making process is as important as the spatial distributions of eye movements. Consequently, new metrics are necessary to effectively assess eye movement data of animations, such as the analysis of meaningful eye movement sequences and eye fixations over time to better understand user decision-making processes with animations.

In psychology research, user affective states are often measured by means of *skin conductance devices,* because of their advantages compared with electroencephalography (EEG) or electromyography (EMG), i.e., they are less obtrusive, and relatively easier to analyse. However, with this methodology, it is difficult to isolate a particular process, e.g., cognitive load, because skin conductance is an indicator of many psycho-physiological processes. It is also difficult to classify arousal signals to the valence of emotions. We attempted to solve this limitation by correlating self-reported stress factors (SSSQ responses) with arousals measured by EDA. However, this method is time-consuming to analyse, as it requires many steps. A future approach would be to use electromyography (EMG), with which it is possible to measure facial expressions directly coupled with positive and negative emotions during a user study.

The *triangulation (or cross-validation) approach* of response data collected from different sources was a promising method to analyse and evaluate users' perceptual, cognitive, and affective processes involved in information extraction and decision-making with animations. One problem might be that the different data sources are recorded at different sampling rates (e.g., eye movements: 300 Hz; EEG: 128 Hz; EDA: 10 Hz). Different sources might also show different signal latency durations, i.e., time between stimulus and event-related response. Event-related potentials (ERPs) of EDA data occur approximately 1–5 s after an event, eye fixations have a duration of about 200–300 ms, saccades approximately 30–80 ms, and ERPs of EEG data (e.g., P300, which is related to decision-making) occur roughly 250–500 ms after a stimulus (Maggi & Fabrikant 2014b). This again requires a time-consuming procedure to process data such that they are comparable and can be meaningfully co-registered. In addition, due to the small number of repeating trials, it was not possible to compare across multiple cases of the same trial, e.g. variations in the user affective state across task difficulty levels. Expertise and detailed knowledge of data processing are also necessary across many disciplines to be able to meaningfully process data. Finally, there is a lack of software able to perform automatic data processing, triangulation analysis and result interpretation.

7.6 Design Guidelines or Recommendations for Animations

From the outcomes of our two experiments, design recommendations for animations may be derived (Table 6). They are structured according to the tested VA dimensions 'animation design' and 'user characteristics', the required tasks (i.e., apprehension of movement changes and prediction of future movement patterns) and the tested factors of the two VA dimensions (i.e., animation design – rate of change, and history path; user characteristics – expertise and familiarity, spatial abilities, motivation, and affective states). The last column in Table 6 suggests empirically-based design recommendations for animations.

Table 6: *Design guidelines derived from the findings of the two user studies according to the VA dimensions 'animation design' and 'user characteristics'.*

ANIMATION DESIGN			
VA dimension	**Task**	**Tested factors**	**Design recommendations for animations**
Animation design	**Detection of movement changes**	Rate of change (or smoothness of the transitions between display scenes) (i.e., semi-static animations, abrupt transitions: 1 scene every 4 seconds, vs continuous animations, smooth transitions: 60 Hz)	To *easily notice slow second-order movement changes* (e.g., accelerations of 0.4 kts/s on a ODS radar display) depicted on dynamic scenes, *semi-static animations seem to be more suited* compared to continuous animations.
			To *easily notice perceptually salient real-time movement patterns* (e.g., faster-moving objects), as well to *effectively apprehend the relative motion between 4 or 8 moving objects* depicted on dynamic scenes, *continuous animations seem to be more suited* compared to semi-static animations.
			To easily notice thematically relevant movement changes that are too smooth (e.g., accelerations of 0.4 kts/s on a ODS radar display), continuous animations seem to be less suited than semi-static animations. *Highlighting thematically relevant movement changes with appropriate visual variables* might be of advantage, especially for novices, to easily identify thematic relevant objects on continuous animations (not empirically tested).
	Prediction of future movement patterns		To *effectively predict near-future movement patterns of two coordinated moving objects* (e.g., two converging aircraft moving at similar or different speeds), *continuous animations seem to be more suited* compared to semi-static animations.

	Detection of movement changes and **prediction of future movement patterns**	History path (or trace)	To easily apprehend (and predict future) movement patterns of 2, 4 and 8 moving objects, *the depiction of history path* (i.e., the depiction of current and past moving objects positions) *seems to be advantageous* than without history path (i.e, the depiction of only the current moving object position, without past positions) *on both semi-static and continuous animations*.
User	**Detection of movementchanges** and **prediction of future movement patterns**	Expertise and familiarity	Expertise in a specific domain might help viewers in processing complex visual information with unfamiliar and novel animation design types, by transferring previous knowledge, even without significant compromise of task effectiveness (if the kind of task and movement patterns are similar with familiar task/displays). Consequently, *the choice of the appropriate animation design to depict movement data influences mainly novices and experts to a lesser extent*.
		Spatial abilities	Having superior spatial skills seems to reduce the cognitive effort required to solve the task and thus positively affects task performance. Consequently, *high spatial skills might be advantageous for viewers when processing information with novel and cognitively demanding animation types*.
		Motivation	Viewers seem to be more motivated to perform visuospatial tasks with continuous animations than with semi-static animations. Consequently, *to enhance viewer motivation, animations of movement data should be designed according to continuous animations* instead of semi-static animations.
		Affective states (i.e., engagement, distress and worry)	Viewer affective states seem to affect task performance and seems to be linked with expertise (or training) as well. Experts and high-trained viewers seem to be more engaged, less distressed, and less worried than novices and low-trained viewers. This positively affects their task performance. In addition, experts and highly trained viewers exhibit higher engagement with continuous animations than with semi-static animations. Consequently, *to design affectively engaging animations, continuous animations seem to be more suited* than semi-static animations, especially for expert viewers.

CHAPTER EIGHT

CONCLUSIONS AND OUTLOOK

8.1 Summary and Scientific Contributions

This thesis aimed to investigate how the three main VA dimensions, i.e., (1) use context and respective task characteristics, (2) animation display design, and (3) user characteristics influence viewer visuospatial decision-making with animations. The main research question leading this thesis was the following:

> **How should animations of real-time movement data be designed considering the task and/or use contexts, and user characteristics?**

For this purpose, we set up two user studies in the context of air traffic control (ATC), in which we asked expert and novice participants to solve two typical tasks related to Situational Awareness (SA). The first experiment investigated how participants apprehended aircraft movement changes (i.e., accelerations of aircraft), and the second experiment how expert participants predicted near-future aircraft movement patterns (i.e., future aircraft positions and distances between aircraft). The tasks, as well as the tested animated displays, were designed according to real-world parameters in the context of ATC.

Our empirical findings show that **the effectiveness and efficiency of decision-making with animations depicting movement data are significantly influenced by the three tested VA dimensions mentioned above (i.e., use context/task, animation design and user characteristics)**. Consequently, to answer our general research question mentioned above in other contexts, animations should be designed considering the following task contexts and user characteristics:

- **Animation design vs task contexts:** Animation design (i.e., semi-static animations vs continuous animations) seemed to affect viewers' visuospatial decision-making depending on the task type. We found that *continuous animations seem to be particularly beneficial to accurately solve tasks involving the effective apprehension of the relative motion between four or eight moving objects* (e.g., identification of aircraft moving faster than another aircraft)*, and the prediction of near-future movement patterns of two converging moving objects*. In contrast, *for tasks involving the prompt detection of second-order movement changes (i.e., accelerations of aircraft), semi-static animations seem to be better suited* for this compared to continuous animations. For this kind of task, we found that continuous animations were more suited to easily notice perceptually salient objects, which stand out more on a continuous changing animated display than on semi-static animations. However, this could be a problem if the perceptual salient objects (e.g., the fastest moving aircraft) are not the thematically relevant ones (e.g., accelerating aircraft). In this case, it could be useful to highlight thematically relevant movement changes with appropriate visual variables to make them more distinguishable from irrelevant ones (not tested empirically). However, as emphasized in the next point, we found that the animation design in apprehension tasks affected the ATC novices more than ATC experts.
- **Animation design vs. user characteristics:** The following user characteristics seemed to influence visuospatial decision-making with animations: the viewers' (1) *ATC expertise*, (2) *spatial abilities*, and (3) *affective state*. First, we found that ATC experts performed the typical ATC tasks more accurately compared to ATC novices. In apprehension tasks, due to their training, ATC experts performed the task with both familiar and novel animation design types equally well, guided by top-down (i.e., cognitive) processes. In contrast, ATC novices seemed to be guided more by perceptually salient features due to their inexperience with the task, resulting in a

worse performance compared to ATC experts, especially for the perceptually demanding continuous animations.

Further, we found that higher-spatial and engaged viewers performed better than lower-spatial and less engaged viewers. Motivation of the viewers might influence positively the efficiency and effectiveness of visuospatial decision-making with animations, and it might also be affected by animation design and expertise, as well. Consequently, the choice of the appropriate animation design to depict movement data influenced mainly novices and less so, experts, in a specific application domain. In addition, viewers with high spatial skills might be advantaged when processing visual information with novel and cognitively demanding animation types. Finally, continuous animations seemed to be particularly suited to enhance viewer motivation and engagement when processing spatio-temporal information with animations.

Furthermore, to effectively test animations, this thesis proposed a novel empirical approach:

- **Methodology:** This thesis proposed a novel empirical approach that combined different data sources (i.e., eye movement data, electroencephalography, electrodermal activity, and questionnaires) to evaluate animations, and to better understand users strategies involved in decision-making with animations. For the analysis of eye movement data, we used new analysis methods such as entropy analysis (Krejtz et al. 2015) that helped us to effectively uncover user strategy patterns involved in visuospatial ATC tasks with animations. Further, the analysis of EEG data and EDA data was useful to measure participant motivation, cognitive load and affective states (e.g., stress) during the experiments, and thus to better explain the reasons of their good or bad task performances.

In general, this thesis contributes to GIScience and cartography in the two following important ways:

- First, in **better understanding how affective, cognitive, and perceptual processes might affect the effectiveness and efficiency of decision-making with spatio-temporal data depicted on animations**.
- Second, in proposing **empirically validated design guidelines for perceptually salient, affectively engaging, and cognitively inspired animations**.

8.2 Outlook

In Section 7.5, *Limitations of the Empirical Methods*, different limitations of our empirical methodology have been listed and can be used as inspiration for future works in GIScience and animation visualization research. In particular, the following future research areas are indicated:

- **Use context/task:** This work has focused on an empirical assessment of animations within a special use case: air traffic control. To be able to *generalize results and animation design guidelines* in a more consistent manner, the three VA dimensions related to decision-making with animated displays (i.e., use context/task, animation design, and user-related factors) should be empirically tested in similar research/application domains and task scenarios.

- **Animation design:** I would like to propose three suggestions for future work concerning animation design:

  (1) I encourage researchers to further test and identify *appropriate visual variables* that could help not only experts, but also novices or people with less training and lower visuospatial skills, in effectively processing spatio-temporal data with animations. Possible future experiments could be conducted on suitable visual variables employed for highlighting task-relevant movement patterns by considering the dynamic nature of animations and the type of movement pattern. For example, Boucheix and Lowe (2010, 2013) propose *progressive path cueing* as a promising signalling method to direct user attention to task-relevant information in animations.

  (2) The additional *inclusion of contextual/causal information* (e.g., reference to landmarks or weather information) in animated displays might improve the detection and prediction of current and future task-relevant information, but only a few empirical studies assessing this topic have been conducted to date (Maggi et al. 2016). Hurter et al. (2014a) developed a method to extract wind parameters to effectively recognize and solve future air traffic conflicts between two or more aircraft trajectories. As proposed by our interviewed experts in ATC, an interesting visualization for ATC radar screens would be the depiction of aircraft trajectories coupled with wind magnitude and direction information (synchronization of two forms of spatiotemporal information), as to better identify aircraft movement changes and predict future movement states. The visualization and comprehension of the causal

relations among moving objects might allow users to better identify and anticipate movement changes (Shipley et al. 2013). For example, this could be performed by comparing static visualizations of future states based on simple models, such as the depiction of a moving object constantly moving at the same velocity and same direction, with continuous animations.

(3) Spatio-temporal actions and interactions between moving objects might be effectively depicted in a more *qualitative manner* rather than quantitatively (Gottfried 2011). For example, consider the situation of imaging two aircraft that are moving in parallel, and then, suddenly, one aircraft changes its orientation approaching the second aircraft. This situation has no significant apparent meaning if the two objects are considered as single independent entities, but it becomes a critical situation if they are considered interdependently. Klippel et al. (2008) highlighted the importance of *topological relations* (according to topological models of static spatial relations), and thus of qualitative parameters, that the participants during their experiments used to characterize and describe relationships between geographic regions moving relative to each other. They probably conceived relationships among moving objects as topological schemata as a simplification and abstraction of the reality. As reported by Gottfried (2011), discovering and visualizing relationships between moving entities is particularly important in such application domains as ecology, surveillance systems, and environmental monitoring or transportation. This is not only important for the users for recognizing and predicting events and actions, but for describing *qualitatively the typology of relative motion between moving objects*, which also defines their *characteristics, behaviour, and the valence of spatio-temporal events*.

- **User characteristics:** Kriz and Hegarty (2007), as well as Lowe (2015), emphasize the need to better understand the interconnection between bottom-up and top-down mechanisms involved in the information extraction related to animations. Animation design guidelines have thus to be not only based on the cognitive aspects of users, but also on their perception. Because of their transitional nature, perceptual (or bottom-up) processes (e.g., those triggered by perceptually salient objects) gain an important role in animations, more so than in static displays. Cartographers should thus *consider both the perceptual and cognitive processes of users in designing animations*, and analyse the influence of bottom-up and top-down processes in a more systematic manner. We found that, beyond perceptual and cognitive processes, the affective state of users

influenced the way in which they explored spatio-temporal data with animations. I thus suggest to more often *integrate psycho-physiological measurements in future empirical studies*, as an additional data source to better understand how users make decisions with animations and with graphic displays in general. These data might not only provide more insight into the strategies that users adopt in dealing with graphic displays in different emotional and cognitive conditions (e.g., in engaged situations, under time pressure, or with monotonous and long tasks causing fatigue), but they might also help in explaining part of the reasons underlying task performance.

- **Methodology:** The results of the *triangulating approach* presented by Maggi & Fabrikant (2014b) have not been shown here in a detailed manner due to time constraints. As a next step, *user decision-making strategies over time* could be more deeply analysed by presenting results relative to eye movements combined with EDA and EEG data. This approach might be particularly suitable if the main interest is to show how cognitive load and motivation of users influence visuospatial decision-making with animations. We believe that this cross-validation approach will also lead the way in new and promising directions in the context of empirical studies concerning decision-making with animations. For example, recorded eye movement sequences, electrodermal responses (i.e., galvanic skin conductance), and brain activity signals (i.e., EEG signals) of the participants can be triangulated to investigate whether the low performance of ATC novices could be due to higher mental workload or perhaps less motivation when solving the task.

Publication Bibliography

Albert, W.; Dixon, E. (2003): Is This What you Expected? The Use of Expectation Measures in Usability Testing. In *Proceedings of Usability Professionals Association 2003 Conference, Scottsdale, AZ, June 2003*.

Andrienko, G.; Andrienko, N.; Bak, P.; Keim, D.; Kisilevich, S.; Wrobel, S. (2011): A Conceptual Framework and Taxonomy of Techniques for Analyzing Movement. In *Journal of Visual Languages & Computing* 22 (3), pp. 213–232. DOI: 10.1016/j.jvlc.2011.02.003.

Andrienko, G.; Andrienko, N.; Demsar, U.; Dransch, D.; Dykes, J.; Fabrikant, S. I. et al. (2010): Space, Time and Visual Analytics. In *International journal of geographical information science* 24 (10), pp. 1577–1600. DOI: 10.1080/13658816.2010.508043.

Andrienko, G.; Andrienko, N.; Jankowski, P.; Keim, D.; Kraak, M.-J.; MacEachren, A.; Wrobel, S. (2007a): Geovisual Analytics for Spatial Decision Support: Setting the Research Agenda. In *International journal of geographical information science* 21 (8), pp. 839–857.

Andrienko, G.; Andrienko, N.; Wrobel, S. (2007b): Visual Analytics Tools for Analysis of Movement Data. ACM SIGKDD Exploration Newsletter 9 (2), pp. 38–46.

Andrienko, N.; Andrienko, G. (2008): Supporting Visual Exploration of Massive Movement Data. In *Proceedings of the Working Conference on Advanced Visual Interfaces, Napoli, Italy, May 28 - 30, 2008*, pp. 474–475.

Andrienko, N.; Andrienko, G. (2011): Spatial Generalization and Aggregation of Massive Movement Data. In *IEEE transactions on visualization and computer graphics* 17 (2), pp. 205–219. DOI: 10.1109/TVCG.2010.44.

Annett, M. (2002): Handedness and Brain Asymmetry: The Right Shift Theory: Psychology Press.

Antonenko, P.; Paas, F.; Grabner, R.; van Gog, T. (2010): Using Electroencephalography to Measure Cognitive Load. In *Educational Psychology Review* 22 (4), pp. 425–438. DOI: 10.1007/s10648-010-9130-y.

Barkley, Russell A. (1991): The Ecological Validity of Laboratory and Analogue Assessment Methods of ADHD symptoms. In *Journal of Abnormal Child Psychology* 19 (2), pp. 149–178. DOI: 10.1007/BF00909976.

Barto, A.; Mirolli, M.; Baldassarre, G. (2013): Novelty or Surprise? In *Frontiers in psychology* 4, p. 907. DOI: 10.3389/fpsyg.2013.00907.

Battersby, S. E.; Goldsberry, K. P. (2010): Considerations in Design of Transition Behaviors for Dynamic Thematic Maps. In *Cartographic Perspectives* (65), pp. 16–32. DOI: 10.14714/CP65.127.

Ben Rebah, M.; Zanin, C. (2011): Rethinking Dynamic Visual Variables: Towards a Framework of Dynamic Semiology. In *Proceedings, GEOVIZ 2011, Hamburg, Germany, March 11, 2011*.

Benedek, M.; Schickel, R. J.; Jauk, E.; Fink, A.; Neubauer, A. C. (2014): Alpha Power Increases in Right Parietal Cortex Reflects Focused Internal Attention. In *Neuropsychologia* 56, pp. 393–400. DOI: 10.1016/j.neuropsychologia.2014.02.010.

Bertin, J. (1967): Sémiologie Graphique. Paris: Mouton et Gauthier- Villars.

Bétrancourt, M.; Tversky, B. (2000): Effect of Computer Animation on Users' Performance: A review. In *Presses Universitaires de France* 43 (4), pp. 311–329.

Blok, C. (2006): Interactive Animation to Visually Explore Time Series of Satellite Imagery. In David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell et al. (Eds.): Visual Information and Information Systems, vol. 3736. Berlin, Heidelberg: Springer Berlin Heidelberg (Lecture Notes in Computer Science), pp. 71–82.

Boag, C.; Neal, A.; Loft, S.; Halford, G. S. (2006): An Analysis of Relational Complexity in an Air Traffic Control Conflict Detection Task. In *Ergonomics* 49 (14), pp. 1508–1526. DOI: 10.1080/00140130600779744.

Boucheix, J.-M.; Lowe, R. K. (2010): An Eye Tracking Comparison of External Pointing Cues and Internal Continuous Cues in Learning with Complex Animations. In *Learning and Instruction* 20 (2), pp. 123–135. DOI: 10.1016/j.learninstruc.2009.02.015.

Boucsein, W. (1992): Electrodermal Activity. Berlin, Germany: Springer.

Boyle, G. J.; Saklofske, D. H.; Matthews, G. (2015): Measures of Personality and Social Psychological Constructs. London, San Diego: Academic Press.

Braithwaite, J. J.; Watson, D. G.; Jones, R.; Rowe, M. (2013): A Guide for Analysing Electrodermal Activity (EDA) & Skin Conductance Responses (SCRs) for Psychological experiments. In *Psychophysiology* 49, pp. 1017–1034.

Briesemeister, B. B.; Tamm, S.; Heine, A.; Jacobs, A. M. (2013): Approach the Good, Withdraw from the Bad—A Review on Frontal Alpha Asymmetry Measures in Applied Psychological Research. In *PSYCH* 04 (03), pp. 261–267.

Brunken, R.; Plass, J.; Leutner, D. (2003): Direct Measurement of Cognitive Load in Multimedia Learning. In *Educational Psychologist* 38 (1), pp. 53–61. DOI: 10.1207/S15326985EP3801_7.

Bunch, R. L.; Lloyd, R. E. (2006): The Cognitive Load of Geographic Information. In *The Professional Geographer* 58 (2), pp. 209–220. DOI: 10.1111/j.1467-9272.2006.00527.x.

Buziek, G.; Dransch, D.; Rase, W.-D. (2000): Dynamische Visualisierung. Grundlagen und Anwendungsbeispiele für kartographische Animationen. Berlin, Heidelberg: Springer Berlin Heidelberg; Imprint: Springer.

Calderon, N. A.; Arias-Hernandez, R.; Fisher, B. (2014): Studying Animation for Real-Time Visual Analytics: A Design Study of Social Media Analytics in Emergency Management. In *Proceeding of System Sciences (HICSS), Hawaii International Conference on System Sciences, Hilton Waikoloa, HI, USA, January 2014*, pp. 1364–1373. DOI: 10.1109/HICSS.2014.176.

Carroll, J. B. (1993): Human Cognitive Abilities. A Survey of Factor-Analytic Studies: Cambridge: Cambridge University Press.

Cavanagh, P.; Alvarez, G. (2005): Tracking Multiple Targets with Multifocal Attention. In *Trends in cognitive sciences* 9 (7), pp. 349–354. DOI: 10.1016/j.tics.2005.05.009.

Cetek, C. (2009): Realistic Speed Change Maneuvers for Air Traffic Conflict Avoidance and their Impact on Aircraft Economics. In *International Journal of Civil Aviation* 1 (1).

Chevalier, F.; Dragicevic, P.; Bezerianos, A.; Fekete, J.-D. (2010): Using Text Animated Transitions to Support Navigation in Document Histories. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, GA, USA, April 2010*, pp. 683–692. DOI: 10.1145/1753326.1753427.

Choi, J.; Silverman, I. (2003): Processes Underlying Sex Differences in Route-Learning Strategies in Children and Adolescents. In *Personality and Individual Differences* 34 (7), pp. 1153–1166. DOI: 10.1016/S0191-8869(02)00105-8.

Clements-Stephens, A. M.; Rimrodt, S. L.; Cutting, L. E. (2009): Developmental Sex Differences in Basic Visuospatial Processing: Differences in Strategy Use? In *Neuroscience letters* 449 (3), pp. 155–160. DOI: 10.1016/j.neulet.2008.10.094.

Çöltekin, A. (2015): Mix Well Before Use: Understanding the Key Ingredients of User Studies. In *ICA Workshop on Envisioning the Future of Cartographic Research at the International Cartographic Conference, Curitiba, Brazil, August 2015*.

Çöltekin, A.; Fabrikant, S. I.; Lacayo, M. (2010): Exploring the Efficiency of Users' Visual Analytics Strategies Based on Sequence Analysis of Eye Movement Recordings. In *International journal of geographical information science* 24 (10), pp. 1559–1575. DOI: 10.1080/13658816.2010.511718.

Costa, G. (2001): Stress Prevention in Air Traffic Control. International Labour Organization (ILO).

Csikszentmihalyi, M. (1990): Flow: The Psychology of Optimal Experience. New York: Harper & Row.

Csikszentmihalyi, M. (2000): Beyond Boredom and Anxiety. San Francisco: Jossey-Bass Publisher.

Davidson, R. J.; Ekman, P.; Saron, C. D.; Senulis, J. A.; Friesen, W. V. (1990): Approach-Withdrawal and Cerebral Asymmetry. Emotional Expression and Brain Physiology I. In *Journal of Personality and Social Psychology* 58 (2), pp. 330–341. DOI: 10.1037/0022-3514.58.2.330.

Davies, C.; Fabrikant, S. I.; Hegarty, M. (2014): Toward Empirically Verified Cartographic Displays. In Robert R. Hoffman, Peter A. Hancock, Mark Scerbo (Eds.): Cambridge Handbook of Applied Perception Research, Volumes 1-2. West Nyack: Cambridge University Press (Cambridge handbooks in psychology), pp. 711–730.

Delorme, A.; Makeig, S. (2004): EEGLAB. An Open Source Toolbox for Analysis of Single-Trial EEG Dynamics Including Independent Component Analysis. In *Journal of Neuroscience Methods* 134 (1), pp. 9–21. DOI: 10.1016/j.jneumeth.2003.10.009.

Diamond, D. M.; Campbell, A. M.; Park, C. R.; Halonen, J.; Zoladz, P. R. (2007): The Temporal Dynamics Model of Emotional Memory Processing: a Synthesis on the Neurobiological Basis of Stress-Induced Amnesia, Flashbulb and Traumatic Memories, and the Yerkes-Dodson Law. In *Neural plasticity* 2007, pp. 1–33. DOI: 10.1155/2007/60803.

DiBiase, D.; MacEachren, A. M.; Krygier, J. B.; Reeves, C. (1992): Animation and the Role of Map Design in Scientific Visualization. In *Cartography and geographic information science* 19 (4), pp. 201–214. DOI: 10.1559/152304092783721295.

Dodge, S. (2015): From Observation to Prediction: The Trajectory of Movement Research in GIScience. In *Onsrud, H. and Kuhn, W., (Eds.), Advancing Geographic Information Science: The Past and Next Twenty Years. Chapter 9, GSDI Association Press*, pp. 123–136.

Dodge, S.; Weibel, R.; Ahearn, S. C.; Buchin, M.; Miller,J. A. (2016): Analysis of Movement Data. In *International journal of geographical information science* 30 (5), pp. 825–834. DOI: 10.1080/13658816.2015.1132424.

Dodge, S.; Weibel, R.; Lautenschütz, A.-K. (2008): Towards a Taxonomy of Movement Patterns. In *Information Visualization* 7 (3-4), pp. 240–252. DOI: 10.1057/palgrave.ivs.9500182.

Dominguez, C.; Vidulich, M.; Vogel, E.; McMillan, G. (1994): Situation Awareness: Papers and Annotated Bibliography.: Armstrong Laboratory, Human System Center.

Driscoll, I.; Hamilton, D. A.; Yeo, R. A.; Brooks, W. M.; Sutherland, R. J. (2005): Virtual Navigation in Humans: The Impact of Age, Sex, and Hormones on Place Learning. In *Hormones and behavior* 47 (3), pp. 326–335. DOI: 10.1016/j.yhbeh.2004.11.013.

Duchowski, A. T. (2007): Eye Tracking Methodology. Theory and Practice. London: Springer.

Durso, F. T.; Bleckley, M. K.; Dattel, A. R. (2006): Does Situation Awareness Add to the Validity of Cognitive Tests? In *Human Factors: The Journal of the Human Factors and Ergonomics Society* 48 (4), pp. 721–733. DOI: 10.1518/001872006779166316.

Ekman, Paul; Friesen, Wallace Verne; Ellsworth, Phoebe (1972): Emotion in the Human Face. Guidelines for Research and an Integration of Findings. New York, Oxford: Pergamon (Pergamon general psychology series, 11).

Ekstrom, R. B.; French, J. W.; Harman, H. H.; Dermen, D. (1976): Manual for Kit of Factor-Referenced Cognitive Tests. Princeton, New Jersey: Educational Testing Service.

Endsley, M. R. (1995): Toward a Theory of Situation Awareness. In *Human Factors: The Journal of the Human Factors and Ergonomics Society* (37), pp. 32–64.

Ersoy, O.; Hurter, C.; Paulovich, F.; Cantareiro, G.; Telea, A. (2011): Skeleton-Based Edge Bundling for Graph Visualization. In *IEEE transactions on visualization and computer graphics* 17 (12), pp. 2364–2373. DOI: 10.1109/TVCG.2011.233.

Fabrikant, S. I. (2005): Towards an Understanding of Geovisualization with Dynamic Displays. In *Proceedings of American Association for Artificial Intelligence (AAAI) 2005 Spring Symposium Series: Reasoning with Mental and External Diagrams: Computational Modeling and Spatial Assistance. Stanford University, Stanford, CA, March 2005*.

Fabrikant, S. I.; Christophe, S.; Papastefanou, G.; Maggi, S. (2012): Emotional Response to Map Design Aesthetics. In *7th International Conference on Geographical Information Science, Columbus, Ohio*, pp. 18–21.

Fabrikant, S. I.; Goldsberry, K. (2005): Thematic Relevance and Perceptual Salience of Dynamic Geovisualization Displays. Proceedings of 22th ICA/ACI International Cartographic Conference, A Coruña, Spain, July 2005.

Fabrikant, S. I.; Hespanha, S. R.; Hegarty, M. (2010): Cognitively Inspired and Perceptually Salient Graphic Displays for Efficient Spatial Inference Making. In *Annals of the Association of American Geographers* 100 (1), pp. 13–29. DOI: 10.1080/00045600903362378.

Fabrikant, S. I.; Lobben, A. (2009): Introduction: Cognitive Issues in Geographic Information Visualization. In *Cartographica: The International Journal for Geographic Information and Geovisualization* 44 (3), pp. 139–143. DOI: 10.3138/carto.44.3.139.

Fabrikant, S. I.; Rebich, S.; Montello, D. R.; Andrienko, G.; Andrienko, N. (2008a): A Visual Analytics Approach to Evaluate Inference Affordance from Animated Map Displays. In *Proceeding of Fifth International Conference on Geographic Information Science: pre-conference workshop on Geospatial Visual Analytics, Park City, Utah, September 2008*, pp. 1–7. DOI: 10.5167/uzh-10198.

Fabrikant, S. I.; Rebich-Hespanha, S.; Andrienko, N.; Andrienko, G.; Montello, D. R. (2008b): Novel Method to Measure Inference Affordance in Static Small-Multiple Map Displays Representing Dynamic Processes. In *The Cartographic Journal* 45 (3), pp. 201–215. DOI: 10.5167/uzh-8402.

Fabrikant, Sara Irina; Rebich-Hespanha, Stacy; Andrienko, Natalia; Andrienko, Gennady; Montello, Daniel R. (2008c): Novel Method to Measure Inference Affordance in Static Small-Multiple Map Displays Representing Dynamic Processes. In *The Cartographic Journal* 45 (3), pp. 201–215. DOI: 10.1179/000870408X311396.

Fairclough, S. H.; Venables, L. (2006): Prediction of Subjective States from Psychophysiology: A Multivariate Approach. In *Biological psychology* 71 (1), pp. 100–110. DOI: 10.1016/j.biopsycho.2005.03.007.

Figner, B.; Murphy, R. O. (2010): Using Skin Conductance in Judgment and Decision Making Research. In *A handbook of process tracing methods for decision research: A critical review and user's guide*, pp. 163–184.

Fink, A.; Benedek, M. (2014): EEG Alpha Power and Creative Ideation. In *Neuroscience & Biobehavioral Reviews* 44, pp. 111–123. DOI: 10.1016/j.neubiorev.2012.12.002.

Fish, C. S. (2015): Cartographic Challenges in Animated Mapping. In *ICA Workshop on Envisioning the Future of Cartographic Research at the International Cartographic Conference, Curitiba, Brazil, August 2015*.

Fraser, K. L.; Ayres, P.; Sweller, J. (2015): Cognitive Load Theory for the Design of Medical Simulations. In *Simulation in healthcare: journal of the Society for Simulation in Healthcare* 10 (5), pp. 295–307. DOI: 10.1097/SIH.0000000000000097.

French, J. W.; Ekstrom, R. B.; Price, L. A. (1963): Manual for Kit of Reference Tests for Cognitive Factor. Princeton, New Jersey: Educational Testing Service.

Gable, P. A.; Harmon-Jones, E. (2008): Approach-Motivated Positive Affect Reduces Breadth of Attention. In *Psychological Science* 19 (5), pp. 476–482. DOI: 10.1111/j.1467-9280.2008.02112.x.

Garlandini, S.; Fabrikant, S. I. (2009): Evaluating the Effectiveness and Efficiency of Visual Variables for Geographic Information Visualization. Proceeding of the 9th International

Conference, COSIT 2009, Aber-Wrac'h, France, September 2009. In *Springer Berlin Heidelberg* 5756, pp. 195–211. DOI: 10.1007/978-3-642-03832-7_12.

Gartner, G. (2012): openemotionmap.org - Emotional Response to Space as an Additional Concept in Cartography. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XXXIX-B4, 2012* XXXIX-B4, pp. 473–476. DOI: 10.5194/isprsarchives-XXXIX-B4-473-2012.

Gaur, V.; Scassellati, B. (2006): Which Motion Features Induce the Perception of Animacy? In *Proceedings of IEEE International Conference for Development and Learning, Bloomington, Indiana, USA, May 31 - June 3, 2006*.

Gobet, F. (2005): Chunking Models of Expertise. Implications for Education. In *Applied Cognitive Psychology* 19 (2), pp. 183–204. DOI: 10.1002/acp.1110.

Goddard, E.; Clifford, C. W. G. (2013): A New Type of Change Blindness: Smooth, Isoluminant Color Changes are Monitored on a Coarse Spatial Scale. In *Journal of Vision* 13 (20). DOI: 10.1167/13.5.20.

Goldsberry, K.; Battersby, S. E. (2009): Issues of Change Detection in Animated Choropleth Maps. In *Cartographica: The International Journal for Geographic Information and Geovisualization* 44 (3), pp. 201–215. DOI: 10.3138/carto.44.3.201.

Gonzalez, C. (1996): Does Animation in User Interfaces Improve Decision Making? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vancouver, BC, Canada, April 13 - 18, 1996*, pp. 27–34. DOI: 10.1145/238386.238396.

Gottfried, B. (2011): Interpreting Motion Events of Pairs of Moving Objects. In *GeoInformatica* 15 (2), pp. 247–271. DOI: 10.1007/s10707-009-0095-2.

Grabner, R. H.; Stern, E.; Neubauer, A. C. (2003): When Intelligence Loses its Impact: Neural Efficiency During Reasoning in a Familiar Area. In *International Journal of Psychophysiology* 49 (2), pp. 89–98.

Grier, R. A. (2015): Situation Awareness in Command and Control. In Robert R. Hoffman, Peter A. Hancock, Mark W. Scerbo, Raja Parasuraman, James L. Szalma (Eds.): The Cambridge Handbook of Applied Perception Research. New York: Cambridge University Press, pp. 891–911.

Griffin, A. L.; McQuoid, J. (2012): At the Intersection of Maps and Emotion: The Challenges of Spatially Representing Experience. In *Kartographische Nachrichten* 06/12, pp. 291–299.

Griffin, Amy L.; MacEachren, Alan M.; Hardisty, Frank; Steiner, Erik; Li, Bonan (2006): A Comparison of Animated Maps with Static Small-Multiple Maps for Visually Identifying Space-

Time Clusters. In *Annals of the Association of American Geographers* 96 (4), pp. 740–753. DOI: 10.1111/j.1467-8306.2006.00514.x.

Gudmundsson, J.; Laube, P.; Wolle, T. (2012): Computational Movement Analysis. In Wolfgang Kresse, David M. Danko (Eds.): Springer Handbook of Geographic Information. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 423–438.

Hägerstrand, T. (1970): What About People in Regional Science? In *Papers of the Regional Science Association* 24 (1), pp. 6–21. DOI: 10.1007/BF01936872.

Harmon-Jones, E.; Gable, P. A.; Price, T. F. (2012): The Influence of Effective States Varying in Motivational Intensity on Cognitive Scope. In *Frontiers in integrative neuroscience* 6 (73), pp. 1–5. DOI: 10.3389/fnint.2012.00073.

Harrower, M. (2004): A Look at the History and Future of Animated Maps. In *Cartographica: The International Journal for Geographic Information and Geovisualization* 39 (3), pp. 33–42. DOI: 10.3138/7MN7-5132-1MW6-4V62.

Harrower, M. (2007): The Cognitive Limits of Animated Maps. In *Cartographica: The International Journal for Geographic Information and Geovisualization* 42 (4), pp. 349–357. DOI: 10.3138/carto.42.4.349.

Harrower, M.; Fabrikant, S. (2008): The Role of Map Animation for Geographic Visualization. In *Geographic Visualization: Concepts, Tools and Applications (eds M. Dodge, M. McDerby and M. Turner), John Wiley & Sons, Ltd, Chichester, UK. DOI:* 10.1002/9780470987643.ch4.

Hegarty, M.; Kriz, S.; Cate, C. (2003): The Roles of Mental Animations and External Animations in Understanding Mechanical Systems. In *Cognition and Instruction* 21 (4), pp. 209–249. DOI: 10.1207/s1532690xci2104_1.

Hegarty, M.; Waller, D. A. (2005): Individual Differences in Spatial Abilities. In Priti Shah, Akira Miyake (Eds.): The Cambridge Handbook of Visuospatial Thinking. Cambridge: Cambridge University Press, pp. 121–169.

Helton, W. S. (2004): Validation of a Short Stress State Questionnaire. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 48 (11), pp. 1238–1242. DOI: 10.1177/154193120404801107.

Henderson, J. M.; Weeks, P. A.; Hollingworth, A. (1999): The Effects of Semantic Consistency on Eye Movements During Complex Scene Viewing. In *Journal of Experimental Psychology: Human Perception and Performance* 25 (1), pp. 210–228. DOI: 10.1037/0096-1523.25.1.210.

Hockenbury, D. H.; Hockenbury, S. E. (2011): Discovering Psychology. New York: Worth; Basingstoke: Palgrave.

Holmqvist, K. (2011): Eye Tracking: A Comprehensive Guide to Methods and Measures. Oxford: Oxford University Press.

Holyoak, M.; Casagrandi, R.; Nathan, R.; Revilla, E.; Spiegel, O. (2008): Trends and Missing Parts in the Study of Movement Ecology. In *Proceedings National Academy of Sciences 2008. DOI:* 10.1073/pnas.0800483105.

Huang, H.; Klettner, S.; Schmidt, M.; Gartner, G.; Leitinger, S.; Wagner, A.; Steinmann, R. (2014): AffectRoute – Considering People's Affective Responses to Environments for Enhancing Route-Planning Services. In *International journal of geographical information science* 28 (12), pp. 2456–2473. DOI: 10.1080/13658816.2014.931585.

Huff, M.; Papenmeier, F.; Zacks, J. M. (2012): Visual Target Detection is Impaired at Event Boundaries. In *Visual Cognition* 20 (7), pp. 848–864. DOI: 10.1080/13506285.2012.705359.

Hurter, C. (2016): Image-Based Visualization: Interactive Multidimensional Data Exploration. San Rafael, California: Morgan & Claypool Publishers (Synthesis lectures on visualization, 4).

Hurter, C.; Alligier, R.; Gianazza, D.; Puechmorel, S.; Andrienko, G.; Andrienko, N. (2014a): Wind Parameters Extraction from Aircraft Trajectories. In *Computers, Environment and Urban Systems* 47, pp. 28–43. DOI: 10.1016/j.compenvurbsys.2014.01.005.

Hurter, C.; Conversy, S. (2008): Towards Characterizing Visualizations. In T. C. Nicholas Graham, Philippe Palanque (Eds.): Interactive Systems. Design, Specification, and Verification, vol. 5136. Berlin, Heidelberg: Springer Berlin Heidelberg (Lecture Notes in Computer Science), pp. 287–293.

Hurter, C.; Conversy, S.; Gianazza, D.; Telea, A. C. (2014b): Interactive Image-Based Information Visualization for Aircraft Trajectory Analysis. In *Transportation Research Part C: Emerging Technologies* 47, pp. 207–227. DOI: 10.1016/j.trc.2014.03.005.

Hurter, C.; Conversy, S.; Kapp, V. (2008): An Infovis Approach to Compare ATC Comets. In *Proceedings of the 3rd International Conference on Research in Air Transportatoin ICRAT, Fairfax, Virginia, USA, June 2008*.

Hurter, C.; Conversy, S.; Vinot, J.-L. (2009a): Temporal Data Visualizations of Air Trafic Controllers. In *Proceeding of the SIGCHI Conference on Human Factors in Computing Systems, ACM, Boston, USA, April 2009*.

Hurter, C.; Tissoires, B.; Conversy, S. (2009b): FromDaDy: Spreading Aircraft Trajectories Across Views to Support Iterative Queries. In *IEEE transactions on visualization and computer graphics* 15 (6), pp. 1017–1024. DOI: 10.1109/TVCG.2009.145.

Imbert, J.-P.; Hodgetts, H. M.; Parise, R.; Vachon, F.; Dehais, F.; Tremblay, S. (2014): Attentional Costs and Failures in Air Traffic Control Notifications. In *Ergonomics* 57 (12), pp. 1817–1832. DOI: 10.1080/00140139.2014.952680.

Irani, P.; Ware, C. (2003): Diagramming Information Structures Using 3D Perceptual Primitives. In *ACM Transactions on Computer-Human Interaction (TOCHI)* 10 (1), pp. 1–19. DOI: 10.1145/606658.606659.

Itti, L.; Koch, C. (2001): Computational Modelling of Visual Attention. In *Nature reviews. Neuroscience* 2 (3), pp. 194–203. DOI: 10.1038/35058500.

Itti, L.; Koch, C.; Niebur, E. (1998): A Model of Saliency-Based Visual Attention for Rapid Scene Analysis. In *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20 (11), pp. 1254–1259. DOI: 10.1109/34.730558.

Jacob, R. J. K.; Karn, K. S. (2003): Eye Tracking in Human-Computer Interaction and Usability Research. In : The Mind's Eye: Elsevier, pp. 573–605.

Kadaba, N.; Irani, P.; Leboe, J. (2007): Visualizing Causal Semantics Using Animations. In *IEEE transactions on visualization and computer graphics* 13 (6), pp. 1254–1261. DOI: 10.1109/TVCG.2007.70528.

Kalyuga, S.; Ayres, P.; Chandler, P.; Sweller, J. (2003): The Expertise Reversal Effect. In *Educational Psychologist* 38 (1), pp. 23–31. DOI: 10.1207/S15326985EP3801_4.

Khacharem, A.; Spanjers, I. A. E.; Zoudji, B.; Kalyuga, S.; Ripoll, H. (2013): Using Segmentation to Support the Learning from Animated Soccer Scenes: An effect of Prior Knowledge. In *Psychology of Sport and Exercise* 14 (2), pp. 154–160.

Klein, T.; van der Zwan, M.; Telea, A. (2014): Dynamic Multiscale Visualization of Flight Data. In *Proceedings of the 9th International Conference on Computer Vision Theory and Applications, VISAPP, January 2014*.

Klettner, S.; Gartner, G. (2012): Modelling Affective Responses to Space.

Klippel, A.; Worboys, M.; Duckham, M. (2008): Identifying Factors of Geographic Event Conceptualisation. In *International journal of geographical information science* 22 (2), pp. 183–204. DOI: 10.1080/13658810701405607.

Knigge, L.; Cope, M. (2006): Grounded Visualization: Integrating the Analysis of Qualitative and Quantitative Data through Grounded Theory and Visualization. In *Environment and Planning A* 38 (11), pp. 2021–2037. DOI: 10.1068/a37327.

Kobayashi, T.; Medina, R. M.; Cova, T. J. (2011): Visualizing Diurnal Population Change in Urban Areas for Emergency Management. In *The Professional Geographer* 63 (1), pp. 113–130. DOI: 10.1080/00330124.2010.533565.

Koffka, K. (1935): Principles of Gestalt Psychology. New York: Harcourt Brace.

Koussoulakou, A.; Kraak, M. J. (1992): Spatia-Temporal Maps and Cartographic Communication. In *The Cartographic Journal* 29 (2), pp. 101–108. DOI: 10.1179/000870492787859745.

Krapp, K. M. (2005): Psychologists and their Theories for Students. Detroit: Thomson Gale.

Krejtz, K.; Duchowski, A. T.; Szmidt T; Krejtz, I.; González Perilli, F.; Pires, A. et al. (2015): Gaze Transition Entropy. In *ACM Transactions on Applied Perception* 13 (1), pp. 1–20. DOI: 10.1145/2834121.

Krejtz, K.; Szmidt, T.; Duchowski, A. T.; Krejtz, I. (2014): Entropy-Based Statistical Analysis of Eye Movement Transitions. In *Proceeding ETRA 2014 Proceedings of the Symposium on Eye Tracking Research and Applications. ACM New York, NY, USA*, pp. 159–166. DOI: 10.1145/2578153.2578176.

Kriglstein, S.; Pohl, M.; Stachl, C. (2012): Animation for Time-oriented Data: An Overview of Empirical Research, pp. 30–35. DOI: 10.1109/IV.2012.16.

Kriz, S.; Hegarty, M. (2007): Top-Down and Bottom-Up Influences on Learning from Animations. In *International Journal of Human-Computer Studies* 65 (11), pp. 911–930. DOI: 10.1016/j.ijhcs.2007.06.005.

Larsen, R. J.; Diener, E. (1992): Promises and Problems with the Circumplex Model of Emotion. In *In: Clark, MS., editor. Review of personality and social psychology. 13. Newbury Park, CA: Sage*, pp. 25–59.

Lasseter, J. (1987): Principles of Traditional Animation Applied to 3D Computer Animation. In *SIGGRAPH Comput. Graph.* 21 (4), pp. 35–44. DOI: 10.1145/37402.37407.

Laube, P. (2014): Computational Movement Analysis. Cham: Springer (SpringerBriefs in Computer Science).

Lautenschütz, A.-K. (2011): Assessing the Relevance of Context for Visualizations of Movement Trajectories. Doctoral dissertation. University of Zurich.

Law, D. J.; Pellegrino, J. W.; Hunt, E. B. (1993): Comparing the Tortoise and the Hare: Gender Differences and Experience in Dynamic Spatial Reasoning Tasks. In *Psychological Science* 4 (1), pp. 35–40. Available online at http://www.jstor.org/stable/40062500.

Lee, P. U.; Klippel, A. (2005): Dynamic Aspects of Spatial Information in Air Traffic Controller Displays. In *AAAI 2005 SPring Symposium Series, Reasoning with Mental and External Diagrams: Computational Modeling and Spatial Assistance, Stanford.*

Lenney, E. (1977): Women's Self-Confidence in Achievement Settings. In *Psychological Bulletin* 84 (1), pp. 1–13. DOI: 10.1037/0033-2909.84.1.1.

Lewis, J. R. (1990): Psychometric Evaluation of an After-Scenario Questionnaire for Computer Usability Studies. In *ACM SIGCHI Bulletin* 23 (1), pp. 78–81. DOI: 10.1145/122672.122692.

Lloyd, R. E.; Bunch, R. L. (2005): Individual Differences in Map Reading Spatial Abilities Using Perceptual and Memory Processes. In *Cartography and geographic information science* 32 (1), pp. 33–46. DOI: 10.1559/1523040053270774.

Lloyd, R. E.; Bunch, R. L. (2008): Explaining Map-reading Performance Efficiency. Gender, Memory, and Geographic Information. In *Cartography and geographic information science* 35 (3), pp. 171–202. DOI: 10.1559/152304008784864677.

Lloyd, R. E.; Bunch, R. L. (2010): Learning Geographic Information from a Map and Text. Learning Environment and Individual Differences. In *Cartographica: The International Journal for Geographic Information and Geovisualization* 45 (3), pp. 169–184. DOI: 10.3138/carto.45.3.169.

Lowe, R. (1999): Extracting Information from an Animation during Complex Visual Learning. In *European Journal of Psychology of Education* 14 (2), pp. 225–244. DOI: 10.1007/BF03172967.

Lowe, R. (2015): Perceptual Learning in the Comprehension of Animations and Animated Diagrams. In Robert R. Hoffman, Peter A. Hancock, Mark W. Scerbo, Raja Parasuraman, James L. Szalma, Robert R. Hoffman et al. (Eds.): The Cambridge Handbook of Applied Perception Research. New York: Cambridge University Press, pp. 692–710.

Lowe, R.; Boucheix, J.-M. (2010): Attention Direction in Static and Animated Diagrams. In Ashok K. Goel, Mateja Jamnik, N. Hari Narayanan (Eds.): Diagrammatic Representation and Inference, vol. 6170. Berlin, Heidelberg: Springer Berlin Heidelberg (Lecture Notes in Computer Science), pp. 250–256.

Lowe, R.; Schnotz, W. (2008): Learning with Animation. Research Implications for Design. Cambridge: Cambridge University Press.

Luck, S. J. (2005): An Introduction to the Event-Related Potential Technique. Cambridge, London: MIT (Cognitive neuroscience).

Lykken, D. T. (1972): Range Correction Applied to Heart Rate and to GSR Data. In *Psychophysiology* 9 (3), pp. 373–379.

MacEachren, A. M. (1995): How Maps Work. Representation, Visualization, and Design. Pbk. ed. New York, NY, USA: The Guilford Press.

Madeleine de Scudéry (1654): Clélie - Histoire Romaine. Paris: A. Courbé.

Maggi, S.; Fabrikant, S. I. (2014a): Embodied Decision Making with Animations. In *Proceedings of the 8th International Conference on Geographic Information Science, Vienna, Austria, September 2014. DOI:* 10.5167/uzh-105653.

Maggi, S.; Fabrikant, S. I. (2014b): Triangulating Eye Movement Data of Animated Displays. In *Proceeding of the 2nd International Workshop on Eye Tracking for Spatial Research, Vienna, Austria, September 2014* 1241, pp. 27–31.

Maggi, S.; Fabrikant, S. I.; Imbert, J.-P.; Hurter, C. (2016): How Do Display Design and User Characteristics Matter in Animations? An Empirical Study with Air Traffic Control Displays 1. In *Cartographica: The International Journal for Geographic Information and Geovisualization* 51 (1), pp. 25–37. DOI: 10.3138/cart.51.1.3176.

Mateeff, S.; Dimitrov, G.; Genova, B.; Likova, L.; Stefanova, M.; Hohnsbein, J. (2000): The Discrimination of Abrupt Changes in Speed and Direction of Visual Motion. In *Vision Research* 40 (4), pp. 409–415. DOI: 10.1016/S0042-6989(99)00185-6.

Matthews, G.; Szalma, J.; Panganiban, A. R.; Neubauer, C.; Warm, J. S. (2013): Profiling Task Stress with the Dundee Stress State Questionnaire. In *Psychology of stress: New research (2013)*, pp. 49–90.

Mayer, R. E.; Hegarty, M.; Mayer, S.; Campbell, J. (2005): When Static Media Promote Active Learning: Annotated Illustrations Versus Narrated Animations in Multimedia Instruction. In *Journal of Experimental Psychology: Applied* 11 (4), pp. 256–265. DOI: 10.1037/1076-898X.11.4.256.

McClung, S. N.; Kang, Z. (2016): Characterization of Visual Scanning Patterns in Air Traffic Control. In *Computational intelligence and neuroscience* 2016, pp. 1–17. DOI: 10.1155/2016/8343842.

Michotte, A. (1963): The Perception of Causality. New York, NY: Basic Books. Translated from the French by T. R. and E. Miles.

Moellering, H. (1976): The Potential Uses of a Computer Animated Film in the Analysis of Geographical Patterns of Traffic Crashes. In *Accident Analysis & Prevention* 8 (4), pp. 215–227. DOI: 10.1016/0001-4575(76)90007-5.

Montello, D. R.; Lovelace, K. L.; Golledge, R. G.; Self, C. M. (1999): Sex-Related Differences and Similarities in Geographic and Environmental Spatial Abilities. In *Annals of the Association of American Geographers* 89 (3), pp. 515–534. DOI: 10.1111/0004-5608.00160.

Morrison, J. L. (1974): A Theoretical Framework for Cartographic Generalization with the Emphasis on the Process of Symbolization. In *International Yearbook of Cartography* 14, pp. 115–127.

Neave, N.; Hamilton, C.; Hutton, L.; Tildesley, N.; Pickering, A. T. (2005): Some Evidence of a Female Advantage in Object Location Memory Using Ecologically Valid Stimuli. In *Human Nature* 16 (2), pp. 146–163. DOI: 10.1007/s12110-005-1001-8.

Neisser, U. (1967): Cognitive Psychology. New York: Meredith.

Newcombe, N. S.; Frick, A. (2010): Early Education for Spatial Intelligence Why, What, and How. In *Mind, Brain, and Education* 4 (3), pp. 102–111. DOI: 10.1111/j.1751-228X.2010.01089.x.

Niessen, C.; Eyferth, K.; Bierwagen, T. (1999): Modelling Cognitive Processes of Experienced Air Traffic Controllers. In *Ergonomics* 42 (11), pp. 1507–1520. DOI: 10.1080/001401399184857.

Nold, C. (2009): Emotional Cartography: Technologies of the Self: Edited by Christian Nold.

Nossum, A. S. (2013): Semistatic Animation – Integrating Past, Present and Future in Map Animations. In *The Cartographic Journal* 49 (1), pp. 43–54. DOI: 10.1179/1743277411Y.0000000014.

Olsen, A. (2012): The Tobii I-VT Fixation Filter. Copyright © Tobii Technology AB.

Papastefanou, G. (2009): Ambulatorisches Assessment: Eine Methode (auch) für die Empirische Sozialforschung. In Martin Weichbold, Johann Bacher, Christof Wolf (Eds.): Umfrageforschung. Herausforderungen und Grenzen. Wiesbaden: VS Verlag für Sozialwissenschaften (Österreichische Zeitschrift für Soziologie. Sonderheft, 9/2009), pp. 443–468.

Pessoa, L. (2009): How do Emotion and Motivation Direct Executive Control? In *Trends in cognitive sciences* 13 (4), pp. 160–166. DOI: 10.1016/j.tics.2009.01.006.

Petersen, H. E.; Dugas, D. J. (1972): The Relative Importance of Contrast and Motion in Visual Target Detection. In *Human Factors: The Journal of the Human Factors and Ergonomics Society* 14, pp. 207–216.

Peuquet, D. J. (2002): Representations of Space and Time. New York, London: Guilford Press.

Plutchik, R. (2001): The Nature of Emotions. In *American Scientist* 89 (4), p. 344. DOI: 10.1511/2001.4.344.

Ponsoda, V.; Scott, D.; Findlay, J. M. (1995): A Probability Vector and Transition Matrix Analysis of Eye Movements During Visual Search. In *Acta Psychologica* 88 (2), pp. 167–185. DOI: 10.1016/0001-6918(95)94012-Y.

Przybylski, A. K.; Rigby, C. S.; Ryan, R. M. (2010): A Motivational Model of Video Game Engagement. In *Review of General Psychology* 14 (2), pp. 154–166. DOI: 10.1037/a0019440.

Pylyshyn, Z. W.; Storm, R. W. (1988): Tracking Multiple Independent Targets: Evidence for a Parallel Tracking Mechanism. In *Spatial Vision* 3 (3), pp. 179–197.

Rantanen, E. M.; Nunes, A. (2005): Hierarchical Conflict Detection in Air Traffic Control. In *The International Journal of Aviation Psychology* 15 (4), pp. 339–362. DOI: 10.1207/s15327108ijap1504_3.

Rieber, L. P. (1991): Animation, Incidental Learning, and Continuing Motivation. In *Journal of Educational Psychology* 83 (3), pp. 318–328. DOI: 10.1037/0022-0663.83.3.318.

Risko, E. F.; Kingstone, A. (2015): Attention in the Wild. In Robert R. Hoffman, Peter A. Hancock, Mark W. Scerbo, Raja Parasuraman, James L. Szalma (Eds.): The Cambridge Handbook of Applied Perception Research. New York: Cambridge University Press, pp. 466–487.

Robertson, G.; Fernandez, R.; Fisher, D.; Lee, B.; Stasko, J. (2008): Effectiveness of Animation in Trend Visualization. In *IEEE transactions on visualization and computer graphics* 14 (6), pp. 1325–1332.

Rohrer, C. (2014): When to Use Which User-Experience Research Methods. Edited by Nielsen Norman Group. Evidence-Based User Experience Research, Training, and Consulting. Available online at https://www.nngroup.com/articles/which-ux-research-methods/, checked on 1/22/2017.

Roseman, I. J. (2014): Motivations and Emotivations: Approach, Avoidance, and Other Tendencies in Motivated and Emotional Behavior. In : Handbook of Approach and Avoidance Motivation: Routledge.

Roth, R. E. (2013): Interactive Maps: What We Know and What We Need to Know. In *Journal of Spatial Information Science* (6). DOI: 10.5311/JOSIS.2013.6.105.

Russell, J. A. (1980): A Circumplex Model of Affect. In *Journal of Personality and Social Psychology* 39, pp. 1161–1178.

Salthouse, T. A. (1990): Working Memory as a Processing Resource in Cognitive Aging. In *Developmental Review* 10 (1), pp. 101–124. DOI: 10.1016/0273-2297(90)90006-P.

Salthouse, T. A. (2009): When Does Age-Related Cognitive Decline Begin? In *Neurobiology of aging* 30 (4), pp. 507–514. DOI: 10.1016/j.neurobiolaging.2008.09.023.

Schlienger, C.; Conversy, S.; Chatty, S.; Anquetil, M.; Mertz, C. (2007): Improving Users' Comprehension of Changes with Animation and Sound: An Empirical Assessment. In *In: Baranauskas C., Palanque P., Abascal J., Barbosa S.D.J. (eds) Human-Computer Interaction –*

INTERACT 2007. INTERACT 2007. Lecture Notes in Computer Science 4662, pp. 207–220. DOI: 10.1007/978-3-540-74796-3_20.

Schnotz, W.; Böckheler, J.; Grzondziel, H. (1999): Individual and Co-Operative Learning with Interactive Animated Pictures. In *European Journal of Psychology of Education* 14 (2), pp. 245–265. DOI: 10.1007/BF03172968.

Shanmugasundaram, M.; Irani, P.; Gutwin, C. (2007): Can Smooth View Transitions Facilitate Perceptual Constancy in Node-Link Diagrams? In *Proceedings of Graphics Interface, ACM New York, NY, USA, Montreal, Canada, May 2007*, pp. 71–78. DOI: 10.1145/1268517.1268531.

Shipley, T. F.; Fabrikant, S. I.; Lautenschütz, A.-K. (2013): Creating Perceptually Salient Animated Displays of Spatiotemporal Coordination in Events. In Martin Raubal, David M. Mark, Andrew U. Frank (Eds.): Cognitive and Linguistic Aspects of Geographic Space. Berlin, Heidelberg: Springer Berlin Heidelberg (Lecture Notes in Geoinformation and Cartography), pp. 259–270.

Shipley, T. F.; Zacks, J. M. (2008): Understanding Events: From Perception to Action. New York: Oxford University Press.

Slingsby, A.; Strachan, J.; Vidale, P.-.L.; Dykes, J.; Wood, J. (2010): Making Hurricane Tracks Accessible. In *Discovery Exhibition, VisWeek 2010*.

Slocum, T.; Sluter, R.; Kessler, F.; Yoder, S. (2004): A Qualitative Evaluation of MapTime, A Program For Exploring Spatiotemporal Point Data. In *Cartographica: The International Journal for Geographic Information and Geovisualization* 39 (3), pp. 43–68. DOI: 10.3138/92T3-T928-8105-88X7.

Slocum, T. A.; Blok C.; Jiang, B.; Koussoulakou, A.; Montello, D. R.; Fuhrmann, S.; and Nicholas R. H. (2001): Cognitive and Usability Issues in Geovisualization. In *Cartography and geographic information science* (28:1), pp. 61–75. DOI: 10.1559/152304001782173998.

Steelman, K. S.; McCarley, J. S.; Wickens, C. D. (2011): Modeling the Control of Attention in Visual Workspaces. In *Human Factors: The Journal of the Human Factors and Ergonomics Society* 53 (2), pp. 142–153. DOI: 10.1177/0018720811404026.

Stein, E. S. (1989): Air Traffic Controller Scanning and Eye Movements in Search of Information - A Literature Review. In *Atlantic City, NJ: FAA Technical Center*.

Štěrba, Z.; Šašinka, Č.; Stachoň, Z.; Kubiček, P.; Tamm, S. (2014): Mixed Research Design in Cartography: A Combination of Qualitative and Quantitative Approaches.

Stewart, J. A. (1982): Perception of Animacy. University of Pennsylvania. Available online at http://repository.upenn.edu/dissertations/AAI8227322.

Sun, T.; Walsh, C. A. (2006): Molecular Approaches to Brain Asymmetry and Handedness. In *Nature Reviews Neuroscience* 7 (8), pp. 655–662. DOI: 10.1038/nrn1930.

Thomas, J. J.; Cook, K. A. (2005): Illuminating the Path: The Research and Development Agenda for Visual Analytics. Los Alamitos: IEEE Computer Society.

Tobler, W. R. (1970): A Computer Movie Simulating Urban Growth in the Detroit Region. In *Economic Geography* 46, pp. 234–240. Available online at http://www.jstor.org/stable/143141.

Tullis, T.; Albert, B. (2008): Measuring the User Experience. Collecting, Analyzing, and Presenting Usability Metrics. Amsterdam, Boston: Morgan Kaufmann (The Morgan Kaufmann series in interactive technologies).

Tversky, B.; Morrison, J. B.; Betrancourt, M. (2002): Animation: Can it Facilitate? In *International Journal of Human-Computer Studies* 57 (4), pp. 247–262. DOI: 10.1006/ijhc.2002.1017.

van Dantzig, S.; Pecher, Diane; Zwaan, Rolf A. (2008): Approach and Avoidance as Action Effects. In *Quarterly journal of experimental psychology (2006)* 61 (9), pp. 1298–1306. DOI: 10.1080/17470210802027987.

van den Bosch, I.; Salimpoor, V. N.; Zatorre, R. J. (2013): Familiarity Mediates the Relationship Between Emotional Arousal and Pleasure During Music Listening. In *Frontiers in human neuroscience*, pp. 7–534. DOI: 10.3389/fnhum.2013.00534.

van Elzakker, C. P.; Griffin, A. L. (2013): Focus on Geoinformation Users: Cognitive and Use/User Issues in Contemporary Cartography. In *GIM International* 27 (8), pp. 20–23.

Vandenberg, S. G.; Kuse, A. R. (1978): Mental Rotations, a Group Test of Three-Dimensional Spatial Visualization. In *Perceptual and motor skills* 47 (2), pp. 599–604. DOI: 10.2466/pms.1978.47.2.599.

Vasiliev, I. (1997): Mapping Time. In *Cartographica: The International Journal for Geographic Information and Geovisualization* 34 (2), pp. 1–51. DOI: 10.3138/D357-234G-2M62-4373.

Vessey, I.; Galletta, D. (1991): Cognitive Fit: An Empirical Study of Information Acquisition. In *Information Systems Research* 2 (1), pp. 63–84. DOI: 10.1287/isre.2.1.63.

Wai, J.; Lubinski, D.; Benbow, C. P. (2009): Spatial Ability for STEM Domains. Aligning over 50 Years of Cumulative Psychological Knowledge Solidifies its Importance. In *Journal of Educational Psychology* 101 (4), pp. 817–835. DOI: 10.1037/a0016127.

Wang, Y.; Cong, W.; Dong, B.; Wu, F.; Hu, M. (2015): Statistical Analysis of Air Traffic Controllers' Eye Movements. In *Air Traffic Management R&D Seminar*.

Ware, Colin (2013): Information visualization. Perception for design. 3rd. ed. Amsterdam: Elsevier.

Webb, E. J.; Campell, D. T.; Schwart, R. D.; Sechrest, L. (1966): Unobtrusive Measures: Nonreactive Research in the Social Sciences. Chicago: Rand McNally.

Weierich, Mariann R.; Wright, Christopher I.; Negreira, Alyson; Dickerson, Brad C.; Barrett, Lisa Feldman (2010): Novelty as a Dimension in the Affective Brain. In *NeuroImage* 49 (3), pp. 2871–2878. DOI: 10.1016/j.neuroimage.2009.09.047.

Weiss, Elisabeth M.; Kemmler, Georg; Deisenhammer, Eberhard A.; Fleischhacker, W.Wolfgang; Delazer, Margarete (2003): Sex differences in cognitive functions. In *Personality and Individual Differences* 35 (4), pp. 863–875. DOI: 10.1016/S0191-8869(02)00288-X.

Wertheimer, M. (1924): Über Gestalttheorie. Vortrag vor der Kant Gesellschaft Berlin, am 17. Dezember 1924. Symposion: Philosophische Zeitschrift für Forschung und Aussprache, 1924, 1, 39-60.

Wilkening, J.; Fabrikant, S. I. (2011): The Effect of Gender and Spatial Abilities on Map Use Preferences and Performance in Road Selection Tasks. Proceedings, 25th International Cartographic Conference, International Cartographic Association, Jul. 3-8, 2011, Paris, France. Available online at http://www.geo.uzh.ch/~sara/pubs/wilkening_fabs_ica2011.pdf.

Wilson, Andrew D.; Golonka, Sabrina (2013): Embodied Cognition is Not What you Think it is. In *Frontiers in psychology* 4, p. 58. DOI: 10.3389/fpsyg.2013.00058.

Wright, Rebecca; Thompson, William L.; Ganis, Giorgio; Newcombe, Nora S.; Kosslyn, Stephen M. (2008): Training generalized spatial skills. In *Psychonomic Bulletin & Review* 15 (4), pp. 763–771. DOI: 10.3758/PBR.15.4.763.

Yerkes, Robert M.; Dodson, John D. (1908): The relation of strength of stimulus to rapidity of habit-formation. In *J. Comp. Neurol. Psychol.* 18 (5), pp. 459–482. DOI: 10.1002/cne.920180503.

Zacks, J. M. (2008): Event Perception. In *Scholarpedia* 3 (10), p. 3837. DOI: 10.4249/scholarpedia.3837.

Zacks, Jeffrey M.; Speer, Nicole K.; Swallow, Khena M.; Braver, Todd S.; Reynolds, Jeremy R. (2007): Event perception: a mind-brain perspective. In *Psychological Bulletin* 133 (2), pp. 273–293. DOI: 10.1037/0033-2909.133.2.273.

Zacks, Jeffrey M.; Tversky, Barbara (2001): Event structure in perception and conception. In *Psychological Bulletin* 127 (1), pp. 3–21. DOI: 10.1037/0033-2909.127.1.3.

Zeile, P.; Höffken, S.; Papastefanou, G. (2009): Mapping People? – The measurement of Physiological Data in City Areas and the Potential Benefit for Urban Planning. REAL CORP 2009: CITIES 3.0 - Smart, Sustainable, Integrative (March 27), pp. 1–12.

Zinser, Otto; Palmer, Debra L.; Miller, Christy R. (2004): Site Distance, Gender, and Knowledge of Geographic Sites. In *Sex Roles* 51 (11-12), pp. 661–686. DOI: 10.1007/s11199-004-0717-y.

Annex 1: Short Stress State Questionnaire

Please indicate how well each word describes how you feel AT THE MOMENT (check the answer from 1 to 5): **1 = Not at all 2 = A little bit 3 = Somewhat 4 = Very much 5 = Extremely**

	1	2	3	4	5
1. I feel dissatisfied.					
2. I feel alert.					
3. I feel depressed.					
4. I feel sad.					
5. I feel active.					
6. I feel impatient.					
7. I feel annoyed.					
8. I feel angry.					
9. I feel irritated.					
10. I feel grouchy.					
11. I am committed to attaining my performance goals.					
12. I want to succeed on the task.					
13. I am motivated to do the task.					
14. I am trying to figure myself out.					
15. I am reflecting about myself.					
16. I am daydreaming about myself.					
17. I feel confident about my abilities.					
18. I feel self-conscious.					
19. I am worried about what other people think of me.					
20. I feel concerned about the impression I am making.					
21. I expect to perform proficiently on this task.					
22. Generally, I feel in control of things.					
23. I thought about how others have done on this task.					
24. I thought about how I would feel if I were told how I					

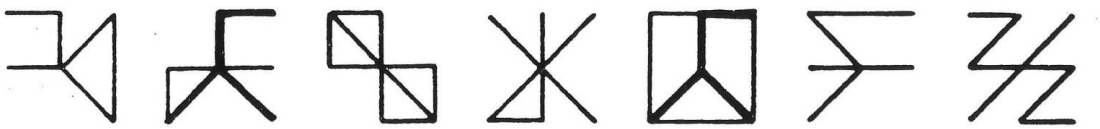
Annex 2: Hidden Pattern Test

How quickly can you recognize a figure that is hidden among other lines?

This test contains many rows of patterns. In each pattern you are to look for the model shown below:

The model must always be in this position, not on its side or upside down.

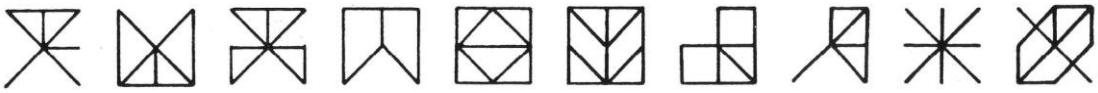
In the next row, when the model appears, it is shown by heavy lines:



(O) (X) (O) (O) (X) (O) (O)

Your task will be to place an X in the space below each pattern in which the model appears and an O below the pattern where the model does not appear. Now, try this row:

1. 2. 3. 4. 5. 6. 7. 8. 9. 10.



( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )

You should have marked an X below patterns 1, 3, 4, 8, and 10, because they contain the model. You should have marked an O below patterns 2, 5, 6, 7, and 9 because they do not contain the model.

Your score on this test will be the number marked correctly minus the number marked incorrectly.

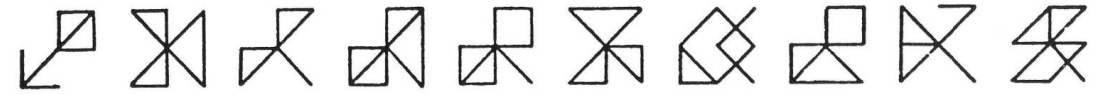
Work as quickly as you can without sacrificing accuracy.

You will have 3 minutes for each of the two parts of this test. Each part has two pages. When you have finished Part I, STOP. Please do not go on to Part II until you are asked to do so.

PLEASE DO NOT TURN THIS PAGE UNTIL ASKED TO DO SO.

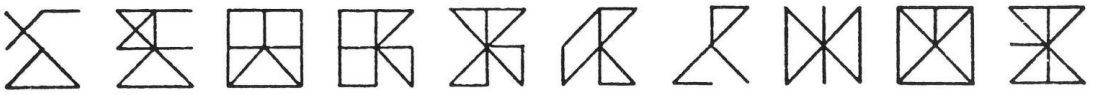
PART I (3 minutes)

Model:

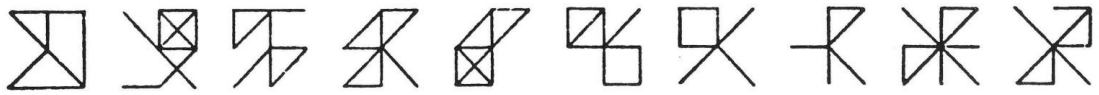


( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )

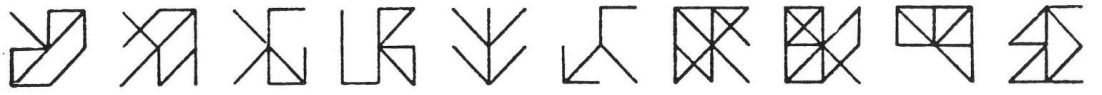
( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )



( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )

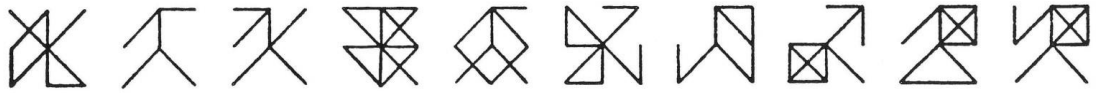


( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )

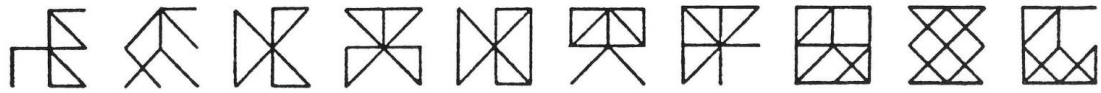


( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )

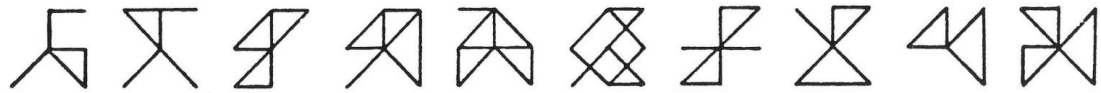
( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )



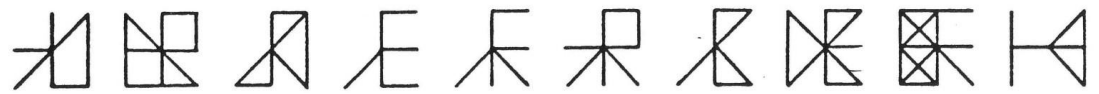
( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )



( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )

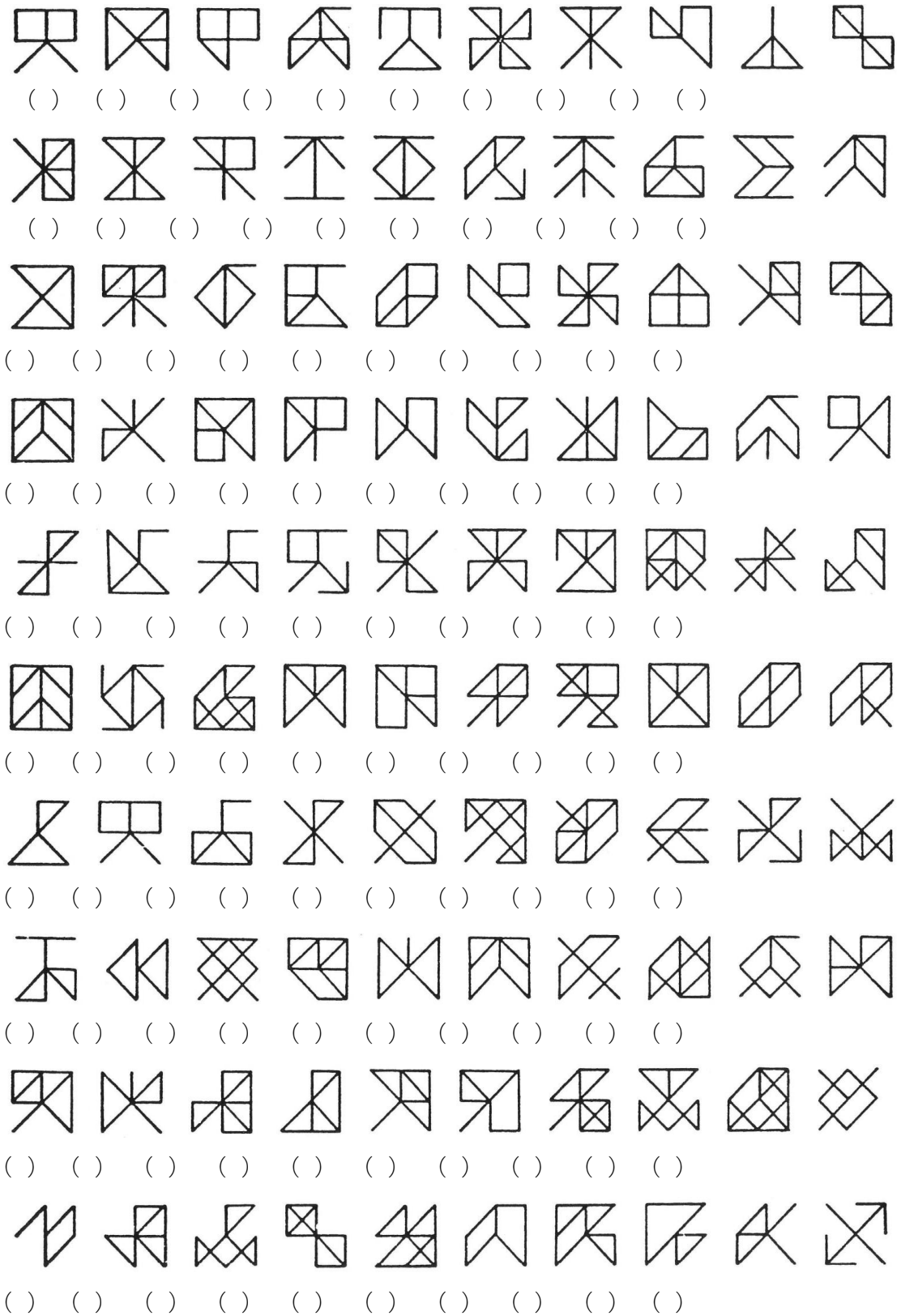


( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )



( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )

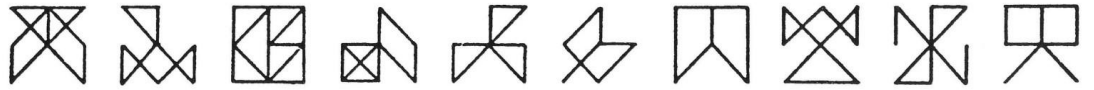
PART I (continued) Model:



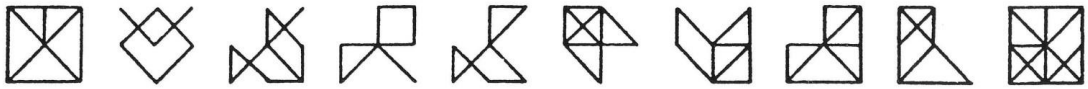
DO NOT GO ON TO THE NEXT PAGE UNTIL ASKED TO DO SO.

PART II (3 minutes)

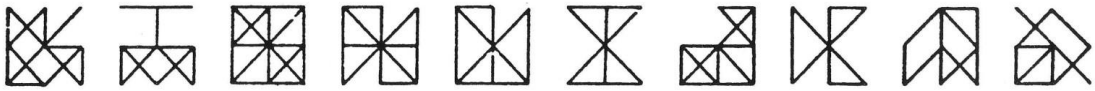
Model:



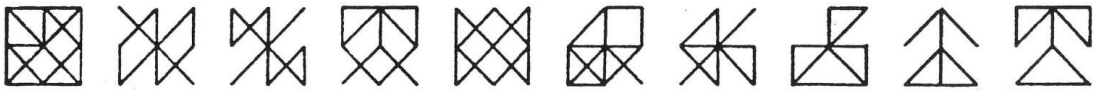
( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )



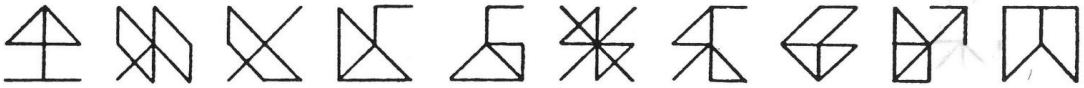
( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )



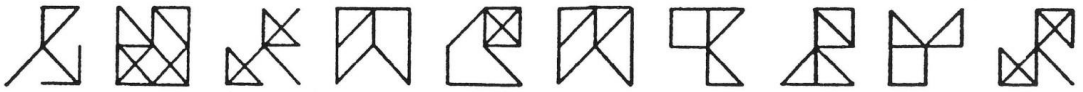
( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )



( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )



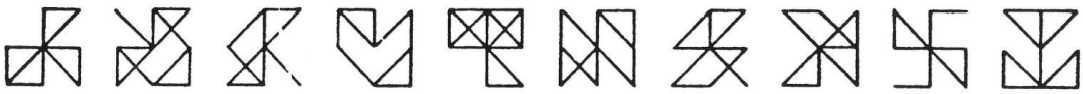
( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )



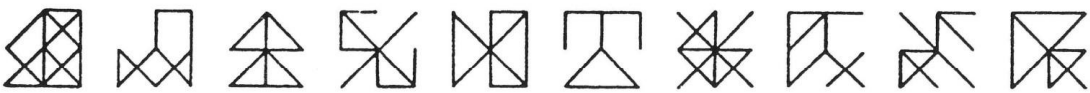
( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )

( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )

( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )



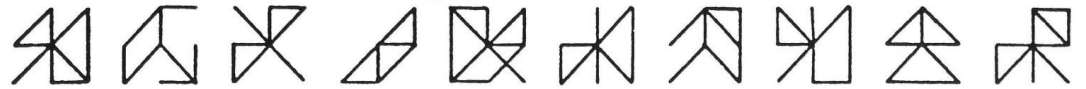
( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )



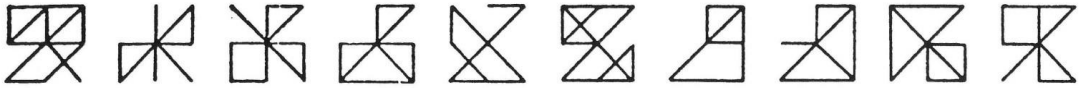
( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )

PART II (continued)

Model:

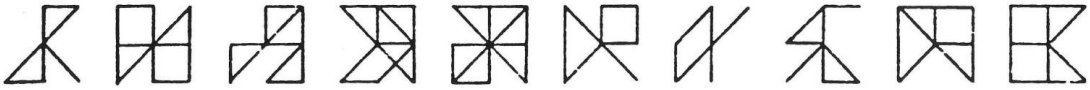


( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )

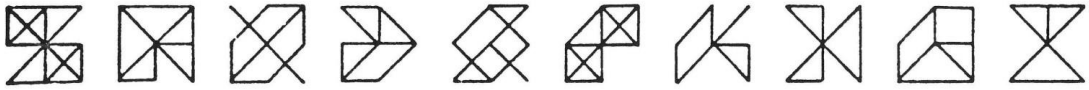


( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )

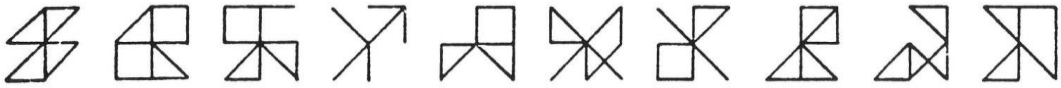
( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )



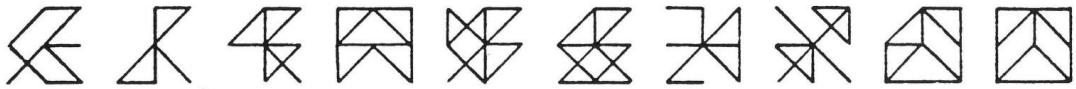
( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )



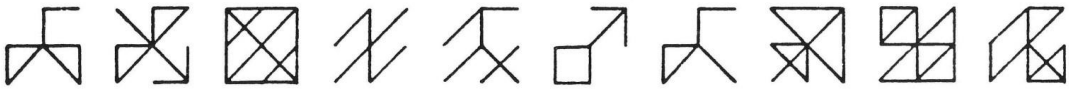
( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )



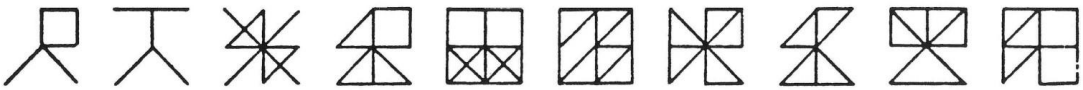
( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )



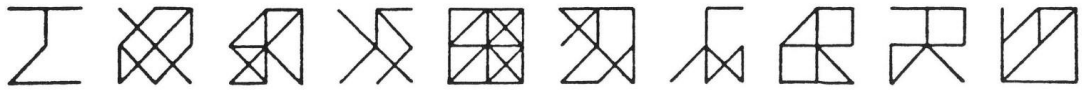
( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )



( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )



( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )



Annex 3: Pre-Test Questionnaire of Experiment I

Please answer to the next questions about **your professional or academic background.**

Your gender:

- ○ Male
- ○ Female

Your age:

[ ]

Are you wearing glasses or contact lenses?

- ○ Yes
- ○ No

Have you been told by a professional that you have imperfect color vision?

- ○ Yes
- ○ No

How would you rate your ability to read maps (e.g., road or topographic maps):

- ○ Poor
- ○ Below average
- ○ Average
- ○ Above average
- ○ Good

How often do you pursue recreational activities that require map reading (for example, hiking, cycling, sailing or orienteering)?

- ○ Never
- ○ Occasionally
- ○ Regularly

How much professional experience, training or college courses have you had in the following domains:

	None	< 1 year	1-2 years	2-5 years	> 5 years
Air Traffic Control	○	○	○	○	○
GIS	○	○	○	○	○
Cartography	○	○	○	○	○
Computer Graphics	○	○	○	○	○
Fine arts, graphic design	○	○	○	○	○

Annex 4: Post-Test Questionnaire of Experiment I

Please answer to the next questions about the **experiment task**, the **display design and parameters.**

Experiment task

Overall, how **easy** was it to solve the experiment task?

Very easy ○ ○ ○ ○ ○ Very difficult

Why?

Overall, how **confident** were you with your answers?

Not confident at all ○ ○ ○ ○ ○ Very confident

Why?

Display Design

Do you use this **kind of visualization** regularly for your work?

Not at all ○ ○ ○ ○ ○ Very often

How **easy** can you detect aircraft movement changes with this kind of display?

Not easy at all ○ ○ ○ ○ ○ Very easy

Why?

How **useful** is this kind of visualization to detect aircraft movement changes?

Not useful at all ○ ○ ○ ○ ○ Very useful

Why?

Did you **enjoy** solving the task with this kind of visualization?

No, not at all ○ ○ ○ ○ ○ Yes, absolutely

Why?

Parameters

How much did you consider the **initial speed** of the aircrafts to detect movement changes:

- ○ Very little or not at all
- ○ Somewhat
- ○ Very much or exclusively

How easily was it to detect movement changes with a **slow initial speed** compared to a **fast initial speed**?

- ○ Easily
- ○ Did not matter
- ○ More difficult

How easily was it to detect movement changes with a **slow acceleration** compared to a **fast acceleration**?

- ○ Easily
- ○ Did not matter
- ○ More difficult

How much did you consider the **distance and/or interactions** between aircraft to detect movement changes?

- ○ Very little or not at all
- ○ Somewhat
- ○ Very much or exclusively

How easily was it to detect movement changes with aircraft in **close distance** to each other compared to aircraft **farther away**?

- ○ Easily
- ○ Did not matter
- ○ More difficult

How much did the **amount of aircraft (4 or 8 aircraft)** have an influence to detect movement changes?

- ○ Very little or not at all
- ○ Somewhat
- ○ Very much or exclusively

How easily was it to detect movement changes with **4 aircraft** compared to **8 aircraft**?

- ○ Easily
- ○ Did not matter
- ○ More difficult

How much did you consider the **distance between the current position and the past positions** of the aircraft (i.e., the 5 squares depicting the aircraft)?

- ○ Very little or not at all
- ○ Somewhat
- ○ Very much or exclusively

Do you have any other **comments** about the experiment?

The experiment and questionnaires are over now. Thank you for participating!

Annex 5: Introduction and Questionnaire for Testing Usability Metrics of Experiment II

You are invited to participate in an experiment, where you are asked to **estimate future aircraft positions** with **four different Air Traffic Control (ATC) display designs**. The aim of this experiment is to find an optimal design to represent aircraft movements.

The design of the tested air traffic control displays is inspired by French radar screens. Standard French radar displays illustrate aircraft as a comet-like form, where each aircraft is represented with five squares of different sizes. The first and largest square represents the current position of the aircraft, followed by gradually smaller squares that illustrate the aircraft's past positions in 2D space.

The tested displays show aircraft positions and movements with **4 different designs**:

- **DESIGN 1:** Current and past aircraft positions updated every 4 seconds (**STANDARD ATC DESIGN**).
- **DESIGN 2:** Current and past aircraft positions updated continuously.
- **DESIGN 3:** Only current aircraft position updated every 4 seconds.
- **DESIGN 4:** Only current aircraft position updated continuously.

DISPLAYS AND TASK

During the experiment, you will see **32 displays** showing **two converging aircraft**. The aircraft are always at the same altitude, but they can move at different speeds.

Your task for this experiment is to **estimate the future positions of the two aircraft where they are clostest to one another**, given the same initial heading and speed.

Each animation will appear for **20 seconds**. After 20 seconds, the animation will stop and you will be asked to indicate the future positions with two mouse clicks.

Before starting with the experiment, please answer the following question:

How well do you expect to be able to estimate future aircraft positions with animated displays, showing…

- **…current and past aircraft positions** updated every **4 seconds:** Not at all ○ ○ ○ ○ ○ Very well
- **…current and past aircraft positions** updated **continuously**: Not at all ○ ○ ○ ○ ○ Very well
- **…only current aircraft position** updated every **4 seconds**: Not at all ○ ○ ○ ○ ○ Very well
- **…only current aircraft position** updated **continuously**: Not at all ○ ○ ○ ○ ○ Very well

Annex 6: Pre-Test Questionnaire of Experiment II

Please answer to the next questions about your personal and professional background.

---

1) Your gender:

- ○ Male
- ○ Female

2) Your age:

[ ]

3) Are you wearing glasses or contact lenses?

- ○ Yes
- ○ No

4) How much professional experience and training have you had in the Air Traffic Control domain:

< 2 years	2 - 4 years	5 -7 years	8 – 10 years	> 10 years
○	○	○	○	○

Annex 7: Post-Test Questionnaire of Experiment II

Please answer the following questions:

---

1) Overall, how **EASY** was it to estimate future aircraft positions with the following animation designs?

   - **DESIGN 1** showing current and past aircraft positions updated every 4 seconds:

     Very easy ○ ○ ○ ○ ○ Very difficult

   - **DESIGN 2** showing current and past aircraft positions updated continuously:

     Very easy ○ ○ ○ ○ ○ Very difficult

   - **DESIGN 3** showing only current aircraft position updated every 4 seconds:

     Very easy ○ ○ ○ ○ ○ Very difficult

   - **DESIGN 4** showing only current aircraft position updated continuously:

     Very easy ○ ○ ○ ○ ○ Very difficult

2) Overall, how **WELL** were you able to estimate future aircraft positions with the following animation designs?

   - **DESIGN 1** showing current and past aircraft positions updated every 4 seconds:

     Not at all ○ ○ ○ ○ ○ Very well

   - **DESIGN 2** showing current and past aircraft positions updated continuously:

     Not at all ○ ○ ○ ○ ○ Very well

   - **DESIGN 3** showing only current aircraft position updated every 4 seconds:

     Not at all ○ ○ ○ ○ ○ Very well

   - **DESIGN 4** showing only current aircraft position updated continuously:

     Not at all ○ ○ ○ ○ ○ Very well

3) What design type **DO YOU PREFER** to estimate future aircraft positions and to visualize aircraft movements?

Please rank the four design types from the **most preferred (1)** to the **last preferred one (4)**:

- **DESIGN 1** showing current and past aircraft positions updated every 4 seconds. ☐
- **DESIGN 2** showing current and past aircraft positions updated continuously. ☐
- **DESIGN 3** showing only current aircraft position updated every 4 seconds. ☐
- **DESIGN 4** showing only current aircraft position updated continuously. ☐

4) In general, what do you think about semi-static and continuous animations to visualize aircraft movements and aircraft dynamics?

5) In general, what do you think about the visualization of the past positions in visualizing aircraft movements and aircraft dynamics?

Annex 8: EDA Analysis

EDA Analysis – Calculation of the area-under-curve (AUC)

Electrodermal activity (EDA) responses of participants were analyzed with the software BIOPAC Acqknowledge (version 4, https://www.biopac.com/product/acqknowledge-software/, last access: 18.01.2017). The online tutorial of BIOPAC can be downloaded here: http://www.biopac.com/wp-content/uploads/EDA-SCR-Analysis.pdf (last access: 18.01.2017).

The steps necessary for our EDA analysis were as follows:

1. *Import the raw data in Acqknowledge.*

   Import raw data with information about the start and end of each stimulus (display) for each participant (**Import → 0.1 sec/sample, none**).

2. *Select the signal within the stimulus and clear the additional signal before and after the experiment.*

   Clear waves that started before the beginning of a stimulus (consider that reaction starts 1-3 seconds after stimulus). At the end of the stimulus, make sure that also the last reaction is inside the considered stimulus interval. Use the "**Clean All**" tool to remove the superfluous waves.

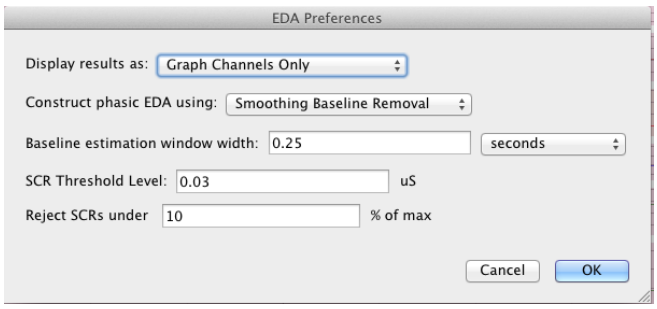
3. *Smooth the raw data (SCLs) to remove noise.*

   Smoothing SCL curve (10, mean), repeat it 5 times (**Transform > Smoothing**). Rename the channel, e.g., SCL.

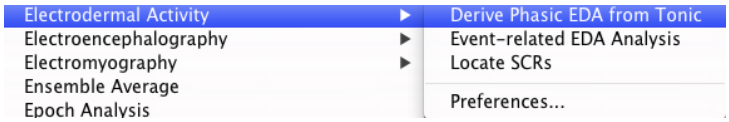
4. *Calculate the SCRs (phasic responses)*

   Derive SCRs with the EDA Analysis tool (**Analysis > Electrodermal Analysis**):

   a. Preferences:



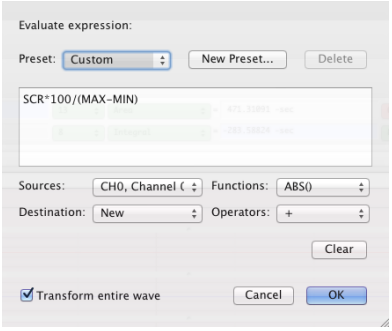
b. Derive SCRs from SCL:



c. Rename the channel, e.g., "SCR".

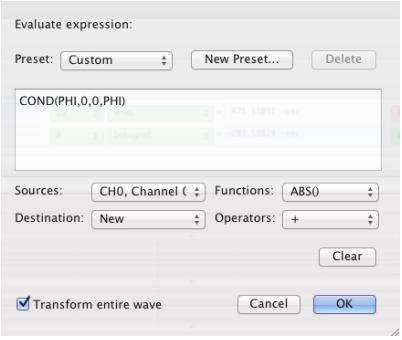
5. *Normalize the SCRs values (PHIs) for each participant (to compare mean values with other participants)*

   Calculate the PHI values with the "Expression" tool (**Transformation > Expression, and write: COND(SCR*100/(MAX-MIN)**). Rename the channel, e.g., PHI.

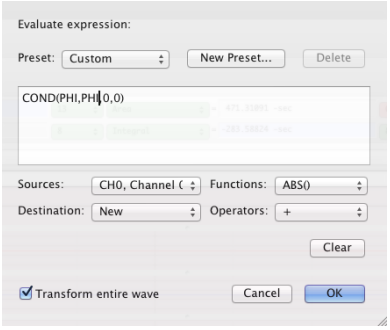


6. *Extract only the positive values of PHIs*

   Extract the positive PHI values with the "Expression" tool (**Transformation > Expression, and write: COND(PHI, 0, 0, PHI)**).



7. *Extract negative PHIs (not so important for the analysis)*



8. *Calculate averaged AUC values*

   Calculate the AUC values, i.e., the integral of the positive PHIs values within the stimulus, with the **Measurement tool**. Then copy the result by means of the right mouse button and "Copy Measurements" on the Measurement bar in an Excel-File or in the Journal.

9. *Additional analyses.*

   It is also possible to calculate the rate of the positive PHIs responses (amount of positive responses / duration) and number of peaks for each stimulus and participant.

10. *Normalize the AUC values with the stimulus duration (only with different stimulus durations among participants).*

    The integral values should be normalized in the case that the stimulus duration has not a fixed value (if the stimulus duration is different for all participants.

11. *Normalize the AUC values with the mean baseline value.*

    Once calculated the AUC value for each stimulus, then normalize it with the averaged baseline value (B) with the following formula:

$$\frac{(\text{Baseline interval EDA integral/s} - \text{Test interval EDA integral/s}) * 100}{(\text{Baseline interval EDA integral/s})}$$

➔ If positive result: decrease of EDA; if negative result: increase of EDA.

Reference paper: Antonenko, P., Paas, F., Grabner, R. and van Gog, T. (2010). Using Electroencephalography (EEG) to measure cognitive load. Educational Psychology Review, 22, pp. 425–438.

EDA Analysis – Statistics

1. *ANOVA of the normalized AUC values.*
2. *AUC values are logarithm-transformed (if they are not normal distributed).*
3. *Inferential statistics with the linear mixed-effects modeling in SPSS (mixed model analysis).*

EDA Analysis –Interpretation of the data

From physiological data, it is possible to derive the overall arousal values by calculating the area-under-curve (AUC), as mentioned above. This gives us the intensity of the arousal value (low/high arousal), but not the valence of the response (positive/negative valence). It is possible to infer the valence of the response by combining the arousal values with others sources, such as the z-changes scores of the Short Stress State Questionnaire (SSSQ) (Helton et al. 2004; Maggi and Fabrikant 2014a; Maggi and Fabrikant 2014b). See also the circumplex model of affect (Russell and Barrett 1999).

Annex 9: EEG Data Analysis

EEG data recorded with Emotiv have been analysed first with the EEGLab tool for Matlab and successively with the STlab tool (2016).

PART I – Procedure with EEGLab

1. **Data import and cleaning**
   - **a.** Launch EEGLab on Matlab
   - **b.** Import EDF data: *File > Import data > Using EEGLAB functions and plugins > From EDF files (BIOSIG toolbox)*
   - **c.** Import event info: *File > Import Event Info > from ASCII- (text) file >, e.g., "latency type", 1, 1, NaN.*
   - **d.** Change channel range (select only the 14 channels from AF3 to AF4): *Edit > Select data > Channel range*
   - **e.** Save and rename your EEGLab file; Locate channels: *Edit > Channel locations (default → ok)*
2. **Filter the data and remove artefacts:**
   - **a.** Filter data with SFIR: *Tools > Filter the data > Short non-linear IIR filterLower: 1Higher: 30*
   - **b.** Manually reject segments with signal artefacts (but not blinks!): *Tool > Reject continuous data by eye > select manually artefacts > REJECT > save file*
   - **c.** Run ICA tool (to remove blinks): *Tool > Run ICA > default*
3. **Reject data using ICA:** *Tools > Reject data using ICA > Reject components by map 1:14; Reject channels (e.g., 2, 3, 8)*
4. **Remove components**
   - **a.** Tools > Remove components > e.g., 2, 3, 8
5. **Create segments that you want to analyse later** (e.g., 4 seconds after stimulus start, and 0.5 seconds before stimulus start for calculate baseline activation): *Tools > Extract epochs; All stimuli*; *From -0.5 to 4 seconds* **>** *ok* **>** *Baseline: -500.0 (-0,5 sec) > ok*
6. **Save datasets according to your test stimuli groups/ independent variables** (e.g., 1 dataset with all stimuli of independent variable A → 8 animations displaying 4 objects, and 1

dataset with all stimuli of independent variable B → 8 animations displaying 8 objects). Save your datasets as e.g.: C01_4_epochs.set and C01_8_epochs.set.

OUTPUT PART I: For each participant two files, "**C(P)#_4_epochs.set**" and "**C(P)#_epochs.set**".

PART II – Procedure with STlab

1. Open Matlab
2. Set the path to the STlab folders
3. Open the directory with the data you want to analyse
4. In Matlab:

   >> faststrafo_eeglab

   Select your files, e.g.: C01_4_epochs.set and C01_8_epochs.set, then OK

   >> vp2group

   Select your files, e.g.: C01_4_epochs.set and C01_8_epochs.set, then OK

   Power[mean(abs(.^2,3))] > OK

   Perform baseline correction? > YES > from e.g., -500 to 0.

   *Save as gr_power_*

   >> gr_areaexport

   Select your file saved before "gr_power_" > OK

   Save as e.g., C01_results > Output: C01_results_data.txt and C01_results_log.txt

5. Calculate FAA scores with the following formula:

   FAA = Ln(R) - ln(L) = ln((F4+FC6+F8)/3) - ln((F3+FC5+F7)/3)

OUTPUT PART II: For each participant one file with FAA scores saved as "**C(P)#_results.xlsx**".

CURRICULUM VITAE

MAGGI Sara Mary Rosa

geboren am 23. August 1980, von Locarno / TI

Ausbildung

Universität Zürich, Zürich

Jan. 2012 – 2017 Promotionsstudium MNF

Dissertation «Depicting Movement Data with Animations for Embodied and Real-Time Decision-Making». Betreuung und Leitung: Prof. Dr. Sara Irina Fabrikant. Promotionskomitee: Prof. Dr. Robert Weibel, Universität Zürich; Prof. Dr. Thomas Shipley, Temple University (USA); Prof. Dr. Christophe Hurter, ENAC (FR).

Sept. 2007 – Nov. 2009 Master of Science in Geographie

Masterarbeit «Evaluation of 3D Spatializations». Betreuung und Leitung: Prof. Dr. Sara Irina Fabrikant und Dr. Arzu Cöltekin, Universität Zürich.

Nebenfach: Physik, Universität Zürich, und Biologie, ETHZ.

Okt. 2003 – Mai. 2006 Bachelor of Science in Geographie

Bachelorarbeit «Mobiles Lernen und Adaptive Lernumgebungen im Bezug zur GIS Ausbildung». Betreuung und Leitung: Prof. Dr. Robert Weibel und Prof. Dr. Dirk Burghardt, Universität Zürich.

Nebenfach: Physik, Universität Zürich, und Biologie, ETHZ.

Okt. 1999 – Mai 2001 1. Vordiplom Biologie, ETHZ.

Liceo Cantonale Locarno, Locarno

Sept. 1996 – Jun. 1999 Eidgenössische Matura mit Schwerpunktfach Geographie.

Publikationen (mit unmittelbarem Bezug zur Dissertation)

Maggi, S., Fabrikant, S.I., Imbert, J.-P., and Hurter, C. (2016). How Do Display Design and User Characteristics Matter in Animations? An Empirical Study with Air Traffic Control Displays. Cartographica. DOI: dx.doi.org/10.3138/cart.3176.

Maggi, S., Fabrikant, S.I., Imbert, J.-P., and Hurter, C. (2015). Quel rôle jouent les principes visuels et les caractéristiques de l'utilisateur dans la visualisation dynamique d'information ? Cartes & Géomatique, Revue du Comité français de cartographie (CFC), traduction de:

„How do Display Design and User Characteristics Matter in Animations? An Empirical Study with Air Traffic Control Displays“, Cartographica (2016).

Maggi, S. and Fabrikant, S.I. (2014). Embodied Decision Making with Animations. Proceedings (extended abstracts), Embodied Decision Making with Animations. 8th International Conference on Geographic Information Science, Sep. 23-26, 2014, Vienna, Austria.

Maggi, S. and Fabrikant, S.I. (2014). Triangulating Eye Movement Data of Animated Displays. CEUR Workshop Proceedings, Vol. 1241, 2nd International Workshop on Eye Tracking for Spatial Research, Sep. 23, 2014, Vienna, Austria: 27-31.

Maggi, S. and Fabrikant, S.I. (2014). Assessing Animated Air Traffic Control Displays. Proceedings (Extended Abstracts), CARTOCON 2014, Olomouc, Czech Republic, Feb. 25-28, 2014.

Maggi, S. and Fabrikant, S.I. (2013). Animated Displays of Moving Objects and Spatiotemporal Coordinated Events. Proceedings (Extended Abstracts), GeoViz Hamburg 2013, Hamburg, Germany. Mar. 6-8, 2013.

Fabrikant, S.I., Christophe, S., Papastefanou, G., and Maggi, S. (2012). Emotional Response to Map Design Aesthetics. Proceedings (Extended Abstracts), GIScience 2012, Columbus, OH, Sep. 18-21, 2012.