

**GAZE INTERACTION:  
COGNITIVE MECHANISMS OF OCULOMOTOR ACTION CONTROL**



**BLICKINTERAKTION: KOGNITIVE MECHANISMEN  
DER OKULOMOTORISCHEN HANDLUNGSKONTROLLE**

DISSERTATION ZUR ERLANGUNG DES NATURWISSENSCHAFTLICHEN DOKTORGRADES  
DER GRADUATE SCHOOL OF LIFE SCIENCES,  
JULIUS-MAXIMILIANS-UNIVERSITÄT WÜRZBURG,  
KLASSE NEUROWISSENSCHAFTEN

VORGELEGT VON  
EVA KATHARINA RIECHELMANN  
AUS FORCHHEIM

WÜRZBURG, 2020

Eingereicht am:

Stempel des GSLS Büros

**Mitglieder des Promotionskomitees:**

Vorsitzende/r: Prof. Dr. Carmen Villmann

1. Betreuer: Prof. Dr. Lynn Huestegge

2. Betreuer: Prof. Dr. Wilfried Kunde

3. Betreuer Prof. Dr. Matthias Gamer

4. Betreuer Prof. Dr. Erhard Wischmeyer

Tag des Promotionskolloquiums: 23.10.2020

Doktorurkunden ausgehändigt am:

## **FINANCIAL SUPPORT**

This research was partly funded by a grant of the German Research Foundation awarded to Lynn Huestegge (HU 1847/7-1).

## **ETHICAL APPROVAL**

All experimental procedures presented in this work involved human subjects and were in accordance with ethical standards (Declaration of Helsinki).

## **CONFLICT OF INTEREST**

The author declares that there are no conflicts of interest.





# LIST OF PUBLICATIONS

Parts of this thesis have been published in the following references:

## Reference [1]

Riechelmann, E., Pieczykolan, A., Horstmann, G., Herwig, A., & Huestegge, L. (2017). Spatio-temporal dynamics of action-effect associations in oculomotor control. *Acta Psychologica*, 180, 130–136. <https://doi.org/10.1016/j.actpsy.2017.09.003>

## Reference [2]

Riechelmann, E., Raettig, T., Böckler, A., & Huestegge, L. (2019). Gaze interaction: Anticipation-based control of the gaze of others. *Psychological Research*. Advance online publication. <https://doi.org/10.1007/s00426-019-01257-4>.

Adapted by permission from Springer Nature Customer Service Centre GmbH.

The publications listed above have been reproduced in part with permission from both the publishers and the co-authors.

A detailed statement of individual author contributions can be found in *Appendix B*.

The following table provides a detailed overview of which parts of the publications were reused where and to which extent within this thesis.

Publication	Usage	Dissertation
<b>Ref. [1]</b>		
p. 1	Text reproduced with minor modifications	pp. 20–21
p. 2	Text reproduced, modified, and extended	pp. 32–34
pp. 2–6	Text reproduced with minor modifications; Figures adapted	pp. 35–48
<b>Ref. [2]</b>		
p. 1	Text reproduced with minor modifications	pp. 1–2
pp. 1–2	Text reproduced, modified, and extended	pp. 3–7
p. 2	Text reproduced, modified, and extended	pp. 19–23
p. 2–3	Text reproduced, modified, and extended	pp. 13–16
p. 3	Text reproduced, modified, and extended	pp. 22–26
p. 4	Text reproduced with minor modifications	pp. 48–50
p. 5–12	Text reproduced with minor modifications; Figures adapted	pp. 51–71
p. 13–17	Text reproduced with minor modifications	pp. 71–81
p. 14	Text reproduced, modified, and extended	pp. 196–198
p. 16	Text reproduced, modified, and extended	pp. 198–199
p. 17	Text reproduced, modified, and extended	pp. 153–154

**The research presented in Chapters 4–6 is currently prepared for publication.**

**Further publications, which are not related to this thesis:**

**Reference [3]**

Riechelmann, E., & Huestegge, L. (2018). Spatial legend compatibility within versus between graphs in multiple graph comprehension. *Attention, Perception & Psychophysics*, *80*(4), 1011–1022. <https://doi.org/10.3758/s13414-018-1484-0>

**Reference [4]**

Riechelmann, E., Weller, L., Huestegge, L., Böckler, A., & Pfister, R. (2019). Revisiting intersubjective action-effect binding: No evidence for social moderators. *Attention, Perception & Psychophysics*, *81*(6), 1991–2002. <https://doi.org/10.3758/s13414-019-01715-6>

**Reference [5]**

Riechelmann, E., Gamer, M., Böckler, A., Huestegge, L. (in press). How ubiquitous is the direct-gaze advantage? Evidence for an averted-gaze advantage in a gaze-discrimination task. *Attention, Perception & Psychophysics*.

**Reference [6]**

Brych, M., Händel, B.F., Riechelmann, E., Pieczykolan, A., & Huestegge, H. (in press). Effects of vocal demands on pupil dilation. *Psychophysiology*.

# ZUSAMMENFASSUNG

Der Mensch nutzt die Augen nicht nur zur Wahrnehmung seiner Umwelt, sondern auch als Handlungsinstrument, um intendierte Effekte in seiner Umwelt zu erzeugen. So werden Blicke beispielsweise dazu verwendet, die Aufmerksamkeit eines anderen auf einen bestimmten Ort zu lenken. Dies weist darauf hin, dass Blickkontrolle einen wichtigen Bestandteil in der sozialen Kommunikation darstellt. Die Forschung zu Blickkontrolle im sozialen Kontext hat sich bisher hauptsächlich auf den *Blick-Empfänger* konzentriert, um die Frage zu beantworten, wie Menschen auf wahrgenommene Blicke reagieren (*Gaze Cueing*). Dieser Ansatz hat dementsprechend bisher kaum den Standpunkt des *Blick-Senders* berücksichtigt. So wurde beispielsweise noch nicht untersucht, welche mentalen Prozesse der *Ausübung* einer Augenbewegung zugrunde liegen, die zum Ziel hat, bei einer anderen Person eine bestimmte Blickreaktion auszulösen. Darüber hinaus werden zielgerichtete Augenbewegungen auch im nicht-sozialen Kontext eingesetzt, beispielsweise beim Entsperren des Smartphones mithilfe der Augen. Diese und andere Beobachtungen zeigen allerdings klar die Notwendigkeit, Blickkontrolle sowohl in der sozialen Kommunikation als auch in anderen, nicht-sozialen Kontexten zu berücksichtigen und dabei gleichzeitig auf Gemeinsamkeiten und Unterschiede zu achten, die der Natur eines sozialen (vs. nicht-sozialen) Handlungskontextes innewohnen. Die vorliegende Arbeit untersucht daher die kognitiven Mechanismen, die solchen zielgerichteten Blickbewegungen in sozialen wie in nicht-sozialen Kontexten zugrunde liegen.

Die in der vorliegenden Arbeit vorgestellten Experimente bauen auf bereits etablierten Paradigmen aus der Forschung zu Okulomotorik und zu basalen kognitiven Prozes-

sen auf. Diese Paradigmen basieren auf dem Prinzip der ideomotorischen Handlungskontrolle, das eine Erklärung für die Entstehung zielgerichteter und beabsichtigter Handlungen liefert. Der ideomotorische Gedanke legt nahe, dass Menschen Assoziationen zwischen ihren Handlungen und den daraus resultierenden Effekten erwerben, die in zwei Richtungen wirken können: Eine Handlung kann die Antizipation ihrer Effekte auslösen, aber die aktive Antizipation eines Handlungseffektes kann auch die damit verbundene Handlung auslösen. Nach der ideomotorischen Theorie beinhaltet Handlungsgenerierung die mentale Antizipation des beabsichtigten Handlungseffektes, um das zugehörige motorische Muster zu aktivieren. Die vorliegenden Experimente beinhalten Situationen, in denen die Probanden den Blick eines virtuellen Gesichts mithilfe ihre eigenen Augenbewegungen steuern. Die im virtuellen Gesicht ausgelösten Blickreaktionen repräsentieren die visuellen Handlungseffekte. Die Situationen werden in Bezug auf die Determinanten von Handlungs-Effekt-Lernen (Kontingenz, Kontiguität, Handlungsmodus während des Lernens) variiert, um die zugrundeliegende Dynamik der okulomotorischen Handlungskontrolle in diesen Situationen zu verstehen. Zusätzlich zu den Gesichtern wurden Handlungseffekte in nicht-sozialen Objekten untersucht, um die Frage zu klären, ob sich die der Blickkontrolle zugrundeliegenden Mechanismen für soziale und nicht-soziale Kontextsituationen unterscheiden.

Die Ergebnisse der vorliegenden Arbeit lassen sich in drei Hauptergebnisse zusammenfassen. 1. Meine Resultate legen nahe, dass Menschen bi-direktionale Assoziationen zwischen ihren Augenbewegungen und der darauf folgenden Blickreaktion einer anderen Person erwerben, was über die Antizipation der beabsichtigten Effekte die okulomotorische Handlungssteuerung beeinflusst. Die beobachteten Ergebnisse zeigen zum ersten Mal, dass Augenbewegungen in einem Blickinteraktionsszenario in Form einer bei der anderen Person ausgelösten Blickreaktion repräsentiert werden. Diese Beobachtung steht im Einklang mit dem ideomotorischen Prinzip der Handlungskontrolle. 2. Die vorliegende Versuchsreihe belegt und erweitert die wegweisenden Ergebnisse von Huestegge und Kreutzfeldt (2012) in

Bezug auf den bedeutenden Einfluss von Handlungseffekten in der okulomotorischen Handlungskontrolle. Ich konnte zeigen, dass sich die Ergebnisse von Huestegge und Kreuzfeldt (2012) über verschiedene Kontexte mit unterschiedlichem Stimulus-Material replizieren lassen unter der Bedingung, dass die wahrgenommenen Handlungseffekte ausreichend stark ausgeprägt waren. 3. Zudem konnte ich zeigen, dass sich Mechanismen der Blickkontrolle in einem sozialen Blickinteraktionskontext vermutlich nicht qualitativ von denen in einem nicht-sozialen Kontext unterscheiden.

Zusammenfassend unterstützen die Ergebnisse die jüngsten theoretischen Überlegungen, die die Rolle von antizipativen Prozessen in der Handlungssteuerung in sozialen Interaktionskontexten betonen. Darüber hinaus legen meine Ergebnisse nahe, dass antizipations-basierte Blickkontrolle im sozialen Kontext auf den gleichen allgemeinen psychologischen Mechanismen wie ideomotorische Blickkontrolle basiert und somit als *integraler* Bestandteil, und nicht als eine spezielle Form der ideomotorischen Blickkontrolle, betrachtet werden sollte.

## SUMMARY

Humans use their eyes not only as visual input devices to perceive the environment, but also as an action tool in order to generate intended effects in their environment. For instance, glances are used to direct someone else's attention to a place of interest, indicating that gaze control is an important part of social communication. Previous research on gaze control in a social context mainly focused on the *gaze recipient* by asking how humans respond to perceived gaze (gaze cueing). So far, this perspective has hardly considered the *actor's* point of view by neglecting to investigate what mental processes are involved when actors decide to perform an eye movement to trigger a gaze response in another person. Furthermore, eye movements are also used to affect the non-social environment, for instance when unlocking the smartphone with the help of the eyes. This and other observations demonstrate the necessity to consider gaze control in contexts other than social communication whilst at the same time focusing on commonalities and differences inherent to the nature of a social (vs. non-social) action context. Thus, the present work explores the cognitive mechanisms that control such goal-oriented eye movements in both social and non-social contexts.

The experiments presented throughout this work are built on pre-established paradigms from both the oculomotor research domain and from basic cognitive psychology. These paradigms are based on the principle of ideomotor action control, which provides an explanatory framework for understanding how goal-oriented, intentional actions come into being. The ideomotor idea suggests that humans acquire associations between their actions

---

and the resulting effects, which can be accessed in a bi-directional manner: Actions can trigger anticipations of their effects, but the anticipated resulting effects can also trigger the associated actions. According to ideomotor theory, action generation involves the mental anticipation of the intended effect (i.e., the action goal) to activate the associated motor pattern. The present experiments involve situations where participants control the gaze of a virtual face via their eye movements. The triggered gaze responses of the virtual face are consistent to the participant's eye movements, representing visual action effects. Experimental situations are varied with respect to determinants of action-effect learning (e.g., contingency, contiguity, action mode during acquisition) in order to unravel the underlying dynamics of oculomotor control in these situations. In addition to faces, conditions involving changes in non-social objects were included to address the question of whether mechanisms underlying gaze control differ for social versus non-social context situations.

The results of the present work can be summarized into three major findings. 1. My data suggest that humans indeed acquire bi-directional associations between their eye movements and the subsequently perceived gaze response of another person, which in turn affect oculomotor action control via the anticipation of the intended effects. The observed results show for the first time that eye movements in a gaze-interaction scenario are represented in terms of their gaze response in others. This observation is in line with the ideomotor theory of action control. 2. The present series of experiments confirms and extends pioneering results of Huestegge and Kreutzfeldt (2012) with respect to the significant influence of action effects in gaze control. I have shown that the results of Huestegge and Kreutzfeldt (2012) can be replicated across different contexts with different stimulus material given that the perceived action effects were sufficiently salient. 3. Furthermore, I could show that mechanisms of gaze control in a social gaze-interaction context do not appear to be qualitatively different from those in a non-social context.

All in all, the results support recent theoretical claims emphasizing the role of anticipation-based action control in social interaction. Moreover, my results suggest that anticipation-based gaze control in a social context is based on the same general psychological mechanisms as ideomotor gaze control, and thus should be considered as an *integral* part rather than as a special form of ideomotor gaze control.



---

# TABLE OF CONTENTS

<b>1</b>	<b>Goal-oriented glances.....</b>	<b>1</b>
<b>2</b>	<b>General introduction.....</b>	<b>3</b>
2.1	Action control .....	3
2.1.1	Effect-based control of action: Ideomotor theory of action control .....	4
2.1.2	Empirical approaches to ideomotor action control.....	6
2.1.3	Sociomotor action control.....	13
2.2	Determinants of action-effect learning.....	16
2.3	Effect-based control of eye movements.....	19
2.3.1	Ideomotor mechanisms in gaze control .....	19
2.3.2	Sociomotor mechanisms in gaze control.....	22
2.4	Overview of the present work .....	26
<b>3</b>	<b>Ideomotor gaze control in social versus non-social contexts .....</b>	<b>31</b>
3.1	Experiment 1 .....	32
3.1.1	Methods.....	35
3.1.2	Results.....	40
3.1.3	Discussion .....	44
3.2	Experiments 2–6.....	48
3.2.1	Methods.....	51
3.2.2	Results.....	59
3.2.3	Discussion .....	71
3.3	Conclusion of Chapter 3.....	81
<b>4</b>	<b>Determinants of action-effect learning in oculomotor action control .....</b>	<b>84</b>
4.1	Effect contingency and effect contiguity in social versus non-social contexts.....	84
4.2	Experiment 7a .....	88
4.2.1	Methods.....	89
4.2.2	Results.....	92
4.2.3	Discussion .....	96

---

4.3	Experiment 7b.....	100
4.3.1	Methods.....	100
4.3.2	Results.....	101
4.3.3	Discussion.....	105
4.4	Experiment 8.....	111
4.4.1	Methods.....	113
4.4.2	Results.....	115
4.4.3	Discussion.....	116
4.5	Conclusion of Chapter 4.....	122
<b>5</b>	<b>The role of the action mode during acquisition.....</b>	<b>124</b>
5.1	Acquisition and effects of action-effect associations in the manual domain.....	124
5.2	Acquisition and effects of action-effect associations in the oculomotor domain.....	127
5.3	Experiment 9.....	130
5.3.1	Methods.....	131
5.3.2	Results.....	138
5.3.3	Discussion.....	141
5.4	Experiment 10.....	146
5.4.1	Methods.....	147
5.4.2	Results.....	148
5.4.3	Discussion.....	149
5.5	Conclusion of Chapter 5.....	150
<b>6</b>	<b>Action-effect compatibility in a gaze leading task.....</b>	<b>153</b>
6.1	Gaze following, joint attention, and gaze leading.....	154
6.2	Experiment 11.....	161
6.2.1	Methods.....	162
6.2.2	Results.....	167
6.2.3	Discussion.....	170
6.3	Experiment 12.....	173
6.3.1	Methods.....	174
6.3.2	Results.....	180
6.3.3	Discussion.....	184
6.4	Conclusion of Chapter 6.....	186
<b>7</b>	<b>General discussion.....</b>	<b>191</b>
7.1	Summary of the results.....	191

---

7.2	Toward a model of shared cognitive mechanisms of oculomotor action control.....	194
7.2.1	Specifications of the model .....	194
7.2.2	Implications on existing oculomotor control theories .....	196
7.3	Open issues and future challenges .....	198
7.3.1	Maintaining experimental control versus generalizability of results .....	198
7.3.2	Further instances of goal-oriented glances.....	200
7.3.3	Gaze control and technical progress .....	203
7.4	Concluding remarks.....	205
	<b>References.....</b>	<b>207</b>
	<b>Appendix.....</b>	<b>XVII</b>
A	Acknowledgements.....	XVII
B	Statement of individual author contributions.....	XXI
C	Affidavit (Eidesstattliche Erklärung).....	XXII



---

# 1 GOAL-ORIENTED GLANCES

Picture yourself at the heart of a party, with an unknown person next to you who attracts your attention. Your goal is to make the person aware of you. One way to attract this person's attention could be to dart a glance at her or him, hoping that your gaze is reciprocated. While you cannot predict another person's reaction for sure, previous experience may provide you with some confidence that others will respond to your approach behavior in the desired way, for example with a shy glance, a smile, or a first "hello" that paves the way for further communication. In fact, daily social interaction teaches us that others often respond quite consistently to our own actions, for example, when we shift our gaze (*gaze cueing*, see *Section 2.3.2*) or when we reach out for a handshake. By anticipating such regularities (i.e., the effects of our own behavior) in human interaction, we are able to plan our actions accordingly and to affect our social environment. However, it is not clear yet which functional mechanisms and representations underlie this type of anticipation-based action control in social interaction. As gaze behavior is a powerful component of social interaction, the present work specifically focuses on eye movements, which aim at affecting the social environment. More specifically, I investigate anticipation-based oculomotor action control by asking how humans control their gaze in order to trigger gaze responses in other human beings. This work is therefore located at the crossroad of several psychological disciplines: It extends research on both oculomotor control, by focusing on the previously neglected aspect of anticipation-based (ideomotor) control mechanisms, and on sociomotor action control, by demonstrating how humans represent and control eye movements that trigger gaze responses in other people. Thus, at the end of this work the gained insights into cognitive

mechanisms of anticipation-based gaze control might even be helpful, for example, to engage with strangers at the next party you will attend more easily.

---

## 2 GENERAL INTRODUCTION

### 2.1 ACTION CONTROL

There is a long-lasting tradition in psychology viewing actions as the outcome of the properties of the preceding input stimuli (see Herwig, Beisert, & Prinz, 2013, for a brief outline of the history of action science). This stimulus-driven view on actions simplifies the complexity of human action in terms of an “input-output” system, thereby neglecting the investigation of motor control, that is, how these actions are brought about. According to Rosenbaum (2005), such research on motor control should actually represent a core question in psychology, when viewing psychology as a scientific discipline which addresses both mind and behavior. In his review, Rosenbaum attributes “the status of a Cinderella in psychological research” (Rosenbaum, 2005, p. 308) to motor control, and gives manifold reasons for this imbalance, ranging from the lacking attractiveness of the field to more fundamental problems of methodological and theoretical nature. Interestingly, already at the end of the nineteenth century, theories about the nature of mechanisms underlying voluntary action control trying to bridge the gap between the input and output side arose. Among others, the ideomotor theory of action control captured the idea that action control is guided by the activated goal of an action, that is, the anticipated (perceptual) effect of this action, instead of merely being determined by external stimulus conditions (Harleß, 1861; Herbart, 1825; James, 1890; Lotze, 1852). However, unfortunately, this ingenious approach to shed light on how humans are able to act in a goal-oriented manner – instead of randomly moving their limbs or passively

waiting for stimulation to respond to – had been forgotten for quite some time, since other psychological disciplines gained popularity, for example, Gestalt psychology or behaviorism (see Herwig et al., 2013). These approaches focused predominantly on perceptual issues, leaving behind any discussions about internal processes that bring about human actions. More than 100 years after the ideomotor idea came up for the first time, however, the theory was revived by Greenwald (1970) and Prinz (1987). Since around that time, researchers from various independent areas of research have shown sustained interest in action control, for example, from the viewpoint of learning, attention, neurophysiology, and neuropsychology, the viewpoint of social and development psychology, or from the perspective of the embodied cognition approach (Herwig et al., 2013). With respect to the contemporary development of ideomotor theory, a considerable number of publications on ideomotor phenomena and recent reformulations of the core ideomotor idea provide evidence for a lasting and growing interest in the investigation of cognitive mechanisms underlying voluntary action control (e.g., see Hommel, 2013, 2019; Hommel, Müsseler, Aschersleben, & Prinz, 2001; Kunde, Weller, & Pfister, 2018; Pfister, 2019; Shin, Proctor, & Capaldi, 2010, for reviews). Since the ideomotor idea also represents the core theoretical framework for the experiments reported in the present work, I will continue by elaborating on the ideomotor idea in general before outlining the most relevant experimental designs used to provide empirical support for ideomotor processes.

### **2.1.1 EFFECT-BASED CONTROL OF ACTION: IDEOMOTOR THEORY OF ACTION CONTROL**

Imagine a typical morning routine: After getting up, you might walk to the bathroom and take a shower. Next, you might switch on the radio in the kitchen and eat some granola. Moreover, you might have a coffee for breakfast. Having this short example episode from everyday life in mind helps to demonstrate two aspects that are important to understand the rationale underlying ideomotor theory. First, coordinated bodily movements are the means



---

for humans to affect their environment, be it by walking toward a location of interest (e.g., bathroom), by pressing a button (e.g., coffee machine), or by grasping for a desired object (e.g., spoon). Second, these bodily movements are usually neither randomly occurring nor are they triggered by external stimulation, but generated in order to perceive an intended change, thus in a goal-oriented manner. More specifically, this means that actions are brought about by translating internal states such as goals and intentions into well-defined muscular activity. Back to the example, this logic would imply that the activated intention to drink a coffee would have the power to initiate pressing the button on the coffee machine.

Ideomotor theory of action control (Harleß, 1861; Herbart, 1825; James, 1890; Lotze, 1852) provides a theoretical framework for understanding how goal-oriented, intentional actions come into being. The theory assumes that learning processes allow people to acquire so-called action-effect associations, that is, associations between motor actions and reliable sensory consequences of these actions through repeated experience of such action-effect couplings. These action-effect associations are assumed to be accessible in a bi-directional manner: Actions can trigger anticipations of their effects, but such anticipations can also trigger their associated actions. In this way, ideomotor theory assumes that, whenever a person pursues the intention to perform a specific action, the activated anticipatory idea of the action's effect, that is the action goal, will guide action selection, initiation, and, finally, action execution (see *Figure 1*). In the case of singing, for example, ideomotor theory assumes that the activated thought of the intended sound triggers the orchestrated interplay of lungs, larynx, vocal folds, and resonating spaces to finally produce the sound (cf. James, 1890). There is notable empirical support for such effect-based action control as described by ideomotor theory (e.g., Dignath, Pfister, Eder, Kiesel, & Kunde, 2014; Elsner & Hommel, 2001; Herwig & Waszak, 2012; Kunde, 2001; Pfeuffer, Kiesel, & Huestegge, 2016; Pfister, Kiesel, & Hoffmann, 2011).

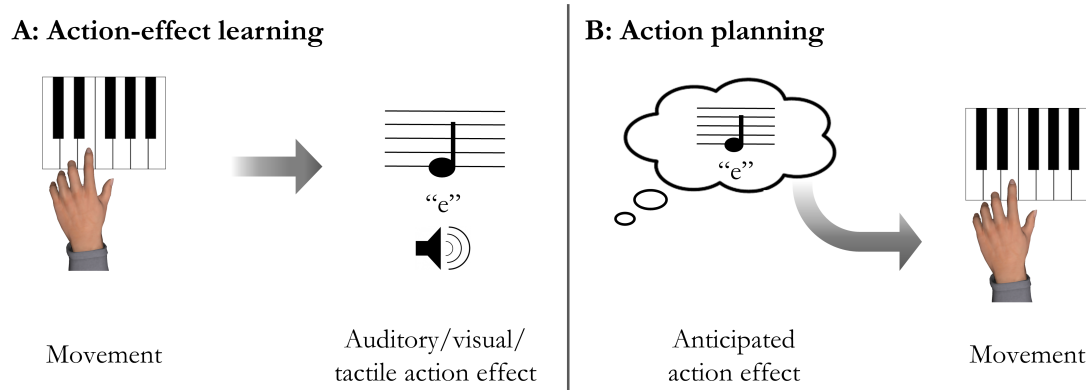


Figure 1. Ideomotor theory proposes that people can acquire associations between movements (e.g., pressing a piano key) and their perceptual effects (A). When anticipating the desired effect, the corresponding motor pattern is activated due to the bi-directional nature of these associations (B).

## 2.1.2 EMPIRICAL APPROACHES TO IDEOMOTOR ACTION CONTROL

Over the years, several empirical approaches have evolved to investigate the mechanisms underlying ideomotor action control. Among them, the most relevant are *action-effect acquisition paradigms* focusing on the acquisition of bi-directional action-effect associations and *action-effect compatibility paradigms* targeting the anticipation of upcoming action effects in action control (see Pfister, 2019, for a recent review). Both empirical approaches yielded a considerable amount of evidence in favor of the basic claims of the ideomotor principle. In the following two paragraphs, I will briefly introduce the two most common empirical paradigms.

### 2.1.2.1 Action-effect acquisition paradigms

Action-effect acquisition paradigms (also called induction paradigms, see Dignath et al., 2014) addressing the acquisition of action-effect associations usually consist of two distinct phases. First, participants repeatedly experience novel action-effect associations in an acquisition phase (e.g., key presses resulting in distinct auditory effects). Thus, an action is contingently followed by a certain sensory effect to internalize the action-effect associations. In a subsequent test phase, the former effect stimuli are presented as imperative stimuli, which are coupled with the same actions produced in the learning phase. Congruency

---

effects – that is, the observation that participants respond more readily with the action that had been associated with the corresponding stimulus in the learning phase as opposed to the action that had been associated with another stimulus during the learning phase are considered empirical markers of action-effect learning (Elsner & Hommel, 2001; Greenwald, 1970).

Among others, the most prominent study applying the two-phase design as described above was from Elsner and Hommel (2001). In their study, participants learned that a left or right key press contingently resulted in high- or low-pitched effect tones during an acquisition phase. Thereby, participants were free to decide on each trial which key to press. For the test phase, Elsner and Hommel relied on two different measures – response time (RT) and choice frequency – to assess whether bi-directional action-effect associations were actually formed during acquisition. Therefore, the authors implemented two kinds of test phases where participants had to respond either in a forced-choice or free-choice mode. The forced-choice test phase required participants to react to the former effect tones with a tone-key mapping that was either congruent or incongruent to the acquisition phase. That is, if a left key press produced a high-pitched tone and a right key press produced a low-pitched tone during acquisition, the presentation of a high-pitched tone as imperative stimulus in the test phase would require a left key press in case of a congruent tone-key mapping, and a right key press in case of an incongruent tone-key mapping. Supporting the ideomotor principle, they observed shorter RTs for an acquisition-congruent mapping compared to an acquisition-incongruent mapping. In the free-choice test phase participants were free to select randomly one of the two responses after a former effect tone and a go-stimulus were presented. Results showed that response frequency of left and right key presses depended on the acquired action-effect association, so that participants preferred to respond to the former effect tone with the key press that had contingently produced this effect in the acquisition phase (Elsner & Hommel, 2001). Their results showed that congruency effects emerged independent of the action mode within the test phase. With respect to the ideomotor principle, these

congruency effects indicate that perceiving a stimulus that served as an action effect in the acquisition phase has the power to activate the associated motor pattern that caused this effect during acquisition via bi-directional action-effect associations. In recent years, several studies using the induction paradigm provided further evidence for the acquisition of bi-directional action-effect associations (e.g., Camus, Hommel, Brunel, & Brouillet, 2018; Dignath et al., 2014; Elsner & Hommel, 2004; Herwig, Prinz, & Waszak, 2007; Hoffmann, Lenhard, Sebald, & Pfister, 2009; Hommel, 2013; Hommel, Alonso, & Fuentes, 2003; Pfister et al., 2011; Wolfensteller & Ruge, 2011).

Similar observations of action priming have been made with experts of a certain domain as participants (e.g., Drost, Rieger, Brass, Gunter, & Prinz, 2005). Experts have usually undergone an extensive training phase in their field. For example, a pianist has experienced countless times that pressing a certain key or a combination of certain keys on the piano always results in a certain auditory effect, namely tones or chords. Thus, expert pianists should have acquired strong action-effect representations. When viewed from an ideomotor perspective, the perception of certain chords should induce the hand movements necessary to produce the corresponding sound due to the bi-directional character of these associations. This is exactly what the study of Drost et al. (2005) has demonstrated. However, while extensive training is sufficient, it is not necessary to observe such effects of action priming. Even the experience of a relatively small number of experimental trials with arbitrary action-effect couplings seems to suffice for the buildup of bi-directional action-effect associations (e.g., Elsner & Hommel, 2001; Hommel, 2009; Wolfensteller & Ruge, 2011). Moreover, some research suggests that action priming occurs even if the action effects are presented at a subliminal level, that is below threshold of conscious perception (Kunde, 2004).

### **2.1.2.2 Action-effect compatibility paradigms**

Studies using the action-effect acquisition paradigm are capable to show that actions and their subsequent effects become associated quickly and that these associations can be

---

activated by perceiving the corresponding effects. However, these studies do not yet lead to the conclusion that anticipation of action effects is pivotal for action generation, which represents the core assumption of the ideomotor theory. Thus, Kunde (2001) developed the action-effect compatibility paradigm (also called response-effect compatibility paradigm)<sup>1</sup> that accounted for this essential claim of the ideomotor idea that goal-oriented actions are selected and initiated by the *anticipation* of future action effects rather than by the *perception* of the same.<sup>2</sup> This experimental approach pursues the logic that the endogenous activation of action effects prior to action generation can only be demonstrated when these sensory action effects are not already perceived in advance of action execution. As a consequence, action effects are exclusively presented *after* action execution in these studies, but the repeated experience of the respective action-effect couplings still allows participants to use the acquired action-effect associations in order to anticipate the upcoming action effects for each of their actions (Pfister, Kiesel, & Melcher, 2010). According to Kunde (2001), the observation of an impact of the upcoming action effects on action selection can be considered as an empirical marker for endogenous effect activation, which is assumed to occur prior to action generation.

On a general level, the studies relying on the action-effect compatibility paradigm manipulate the relation between features of the to-be-executed action (i.e., response) and features of the resulting action effects (Pfister, 2019). Thus, participants execute actions varying on a particular dimension (e.g., intensity of a key press), and the resulting action effects do either overlap with respect to the features of the particular dimension (corresponding

---

<sup>1</sup> The terminology action-effect compatibility paradigm was introduced by Pfister (2019) in order to account for the fact that the responses as performed by the participants in the response-effect compatibility paradigm represented more than responses toward an imperative stimulus only, but rather the means to produce additional effects in the environment, thus actions. I will therefore use the term action throughout this work.

<sup>2</sup> Note that the latter would even be dysfunctional when viewed in conjunction with the fact that perceiving the action effect of a preceding action would result in the activation of the same motor action again, inducing an iterative process without any stopping rule. Please see Greenwald (1970) and Kunde (2001) for a more elaborated discussion.

action-effect mapping, e.g., a forceful key press triggering a loud effect tone) or not (non-corresponding action-effect mapping, e.g., a forceful key press triggering a quiet effect tone). This action (intensity) – effect (loudness) coupling can also be observed in an applied scenario, for example, when playing the piano: While a forceful key press on the piano usually triggers a loud tone, a soft key press would trigger a quiet tone. The action-effect compatibility paradigm follows the rationale of the principle of dimensional overlap, which has action generation shown to be influenced by the compatibility between features of the stimulus and features of the to-be-executed action (stimulus-response compatibility effects; cf. Kornblum, Hasbroucq, & Osman, 1990; Proctor & Reeve, 1990). More specifically, when stimulus and to-be-executed action overlap with respect to one dimension (e.g., spatial), the features activated by the presentation of the stimulus facilitate to-be-executed actions with compatible features (Fitts & Seeger, 1953). For example, it is easier to respond to a stimulus presented on the left side of the screen with a left key press as compared to a right key press, even though the location of the stimulus is task-irrelevant (also known as Simon-Effect, cf. Simon, 1990). Thus, Kunde (2001) argued that if action planning actually entails the mental representation of the intended action effect prior to action execution as proposed by ideomotor theory, the mere mental representation should have the same or similar power to prime the features of the anticipated effects and thereby affect action execution, just like the actual perception of these effect stimuli would do in stimulus-response compatibility experiments (cf. Tlauka & McKenna, 1998, for evidence for stimulus-response compatibility effects in imagery). Based on these considerations, the ideomotor principle predicts the same compatibility effects to come into effect between to-be-executed actions and their effects as has been demonstrated between stimuli and responses.

This is precisely what the study of Kunde (2001) has demonstrated: It is easier to generate motor actions which consistently produce action effects that resemble rather than differ from the executed action itself with respect to spatial features. More specifically, the

---

experimental protocol of Experiment 1 in the study of Kunde (2001) required participants to respond to a non-spatial stimulus attribute (i.e., color) with certain key presses, which consistently resulted in visual effects. The crucial manipulation was that the relation between the location of the responses and the location of their visual effects (i.e., the lighting of four horizontally aligned boxes on the screen). For the corresponding action-effect mapping, each key press resulted in the lighting of the spatially corresponding box, while the box two positions adjacent to the key press was lighted up in case of a non-corresponding action-effect mapping. In line with the expectations, response latencies were reduced for the corresponding as compared to the non-corresponding mapping, even though the action effects appeared exclusively after action execution and were task-irrelevant. Moreover, merely imagined action effects have the power to affect the time necessary to initiate an action within the action-effect compatibility paradigm in a similar way as action effects that are actually perceived (Pfister, Pfeuffer, & Kunde, 2014). The typical influence of action-effect compatibility has been replicated several times for the spatial dimension (e.g., Ansorge, 2002; Pfister, Janczyk, Gressmann, Fournier, & Kunde, 2014; Pfister & Kunde, 2013). Further evidence has been obtained from several different domains of motor control in the past few years. For example, RTs are reduced, when pushing a button forcefully results in the sound of a loud rather than a soft tone (Kunde, 2001, Experiment 2; Kunde, Koch, & Hoffmann, 2004; Paelecke & Kunde, 2007) or when performing a key press with long duration produces an effect tone of long rather than of short duration (Kunde, 2003). Similar observations have, for example, also been made in the domain of speech production (Badets, Koch, & Toussaint, 2013; Koch & Kunde, 2002), musical performance (Keller & Koch, 2008), typing (Rieger, 2007), grasping (Camus et al., 2018), or human-computer interaction (Chen & Proctor, 2013). Similarly, body-related action effects, such as tactile stimulation, have been shown to affect action control in a similar manner as action effects in the environment (Pfister, Janczyk et al., 2014;

Wirth, Pfister, Brandes, & Kunde, 2016), demonstrating the generality of the action-effect compatibility phenomenon.

Most of the studies relying on the action-effect compatibility paradigm utilized a blocked design such that participants first experienced the compatible action-effect mapping, and afterwards the incompatible action-effect mapping (or vice versa). This design was first employed by Kunde (2001), who justified the approach with methodical limitations. He argued that only with the blocked design, where each action consistently triggers a certain effect, participants would be able to rely on these action effects as mental cues to guide action generation. However, as outlined by Pfister et al. (2010), this approach is not compatible with what humans experience in day-to-day situations where the compatibility relations between actions and effects can vary rapidly, depending on the context in which the action is carried out. Pfister et al. (2010) substantiate their argumentation with an example from text production using a word processing program. For example, key presses performed by the left hand on left-hand keys on the computer keyboard produce spatially compatible action effects on the screen when at the beginning of a new line. For instance, pressing the “A” key on a QWERTZ keyboard at the beginning of a new line results in the appearance of the letter “A” on the left side of the screen. However, the same key press produces spatially incompatible action effects when approaching the middle and especially the end of the line (e.g., letter “A” appearing on the right side of the monitor). Thus, if voluntary actions are generated via the recollection of their sensory consequences, as proposed by ideomotor theory, an influence of action-effect compatibility should also be present when action-effect relations vary frequently given the predictability of the effects due to the action context. To account for this variability in action-effect relations, Pfister et al. (2010) adapted the action-effect compatibility design such that action-effect relations varied on a trial-by-trial basis. In line with studies showing that stimulus-response compatibility effects can even be observed with varying stimulus-response relations when the current mapping is pre-cued (Vu & Proctor,



---

2004), the authors observed action-effect compatibility effects for rapidly varying action-effect relations, validating the ideomotor principle as an explanatory account for context-dependent action control.

### 2.1.3 SOCIOMOTOR ACTION CONTROL

Research on action-effect associations has mainly studied the impact of action effects in the inanimate environment, using arbitrary action effects such as tones (e.g., Dignath et al., 2014) or visual effects (e.g., Badets et al., 2013) as the consequence of an individual's action. However, humans are rarely isolated in inanimate environments but interact regularly with other people. In such animate environments, actions have direct consequences for other humans – i.e., those actions affect other people's behavior – or might even be deliberately executed to influence others (Kunde et al., 2018). Moreover, even though one can never safely predict a certain response of another person, there are many situations where actions affect the behavior of another person in a predictable way. To give an everyday example, people use their fingertip to guide the attention of an interaction partner toward an interesting object by pointing toward that object (Herbort & Kunde, 2016, 2018). Observing such a pointing gesture will almost automatically prompt the other person to look at the object as indicated by the fingertip. Recent approaches to understand and incorporate such social consequences of an action into theories of action control resulted in an extension of the ideomotor view to the social domain, namely the sociomotor framework (Kunde et al., 2018). The authors reasoned that agents' actions can elicit predictable but probably less immediate behavioral responses in other humans (i.e., social action effects), which leads to the buildup of intersubjective action-response associations. The theory claims that agents anticipate responses they evoke in others to initiate the associated motor action. Thus, similar processes as involved in ideomotor action control are assumed to underlie sociomotor action control.

However, social processing is assumed to differ from processing of other kinds of information in that it is usually more probabilistic, inexact and more ambiguous (Heider, 1977; Mitchell, 2009) and based on relatively implicit and pre-reflective processes (Frith & Frith, 2008; Vogeley & Roepstorff, 2009). For example, knocking on a door usually results in an immediate (auditory) effect, while social responses typically take more time (lower contiguity) and are less predictable (lower contingency). Consequently, these peculiarities must be taken into account for the investigation of social actions (see Kunde et al., 2018).

There is already a lot of research on how the observation of human actions exerts an influence on the observer (e.g. Becchio, Manera, Sartori, Cavallo, & Castiello, 2012; Herbort, Koning, van Uem, & Meulenbroek, 2012; Heyes, 2011; Hudson, Nicholson, Simpson, Ellis, & Bach, 2015; Shiffrar & Freyd, 1990). It is known, for example, that the mere observation of another person's movement enables humans to attribute a specific intention to this person (Becchio et al., 2012), and that humans tend to spontaneously imitate such observed actions (Heyes, 2011). Most of these studies, however, consider actions of other people (e.g., bodily movements, gaze behavior etc.) as stimuli that serve to trigger behavioral responses in others, thereby focusing on human responses to observed social stimuli or actions. The sociomotor account, however, pursues a different strategy: It is the focus on cognitive processes on the actor's side that assigns a unique role to sociomotor action control within the field of social cognition. That is, in a standard social situation, where person A performs an action (e.g., pointing toward the sky) which exerts an influence on the observing person B (e.g., looking up at the sky), the sociomotor account emphasizes the role of the acting person A in such a situation. Relatedly, the recent request to incorporate more interactive study designs and to investigate social interaction from the viewpoint of the actor has also been raised by Schilbach (2014) for the field of social neuroscience in general, and by Pfeiffer, Vogeley, and Schilbach (2013) for the study of gaze in social interaction in particular.

---

Evidence for the acquisition of bi-directional associations between one's own actions and responses of others to these actions (so-called social action effects) comes, for example, from Sato and Itakura (2013), who observed evidence for learned associations between key presses and the resulting mouth gesture of a target face. Interestingly, these effects were only present in case of direct eye contact (i.e., when the target's eyes were open compared to averted/closed eyes; but see Riechelmann, Weller, Huestegge, Böckler, & Pfister, 2019, for diverging results). In a similar vein, the acquisition of links between finger movements and imitative (or counter-imitative) finger movements has been shown to influence imitation behavior (Bunlon, Marshall, Quandt, & Bouquet, 2015). Together, these studies provide evidence in favor of the sociomotor approach by demonstrating that social action effects can be bound to preceding motor actions.

There is also support for the sociomotor assumption that anticipated social action effects play a role when generating motor actions causing responses in others. The anticipation of such social action effects have, for example, been studied in the context of imitation, demonstrating that a foreseeably imitative response of a virtual avatar to one's own action (rather than a counter-imitative response) facilitates motor actions (Pfister, Dignath, Hommel, & Kunde, 2013; Pfister, Weller, Dignath, & Kunde, 2017). These results suggest that the anticipated social consequences of one's own action are key to effect-based action. The study of Kunde, Lozo, and Neumann (2011) adapted the action-effect compatibility paradigm (see *Section 2.1.2.2*) such that facial expressions served as both action and action effects. In their study, participants produced smiling or frowning faces on a screen by either contracting their zygomaticus major (a muscle predominantly involved in smiling) or the corrugator supercilii (a muscle predominantly involved in frowning). Participants were faster to produce the facial expressions with a compatible mapping (e.g., the contraction of the zygomaticus major produced a smiling face) than with an incompatible mapping (e.g., the contraction of the zygomaticus major produced a frowning face). The results imply that facial

muscle control is sensitive to the visual effects these muscle contractions produce in others. All in all, these findings highlight the importance of social consequences for human action control.

## 2.2 DETERMINANTS OF ACTION-EFFECT LEARNING

Most of the research on effect-based action control cited so far has in common that these studies relied on a “perfect” world where the execution of a motor action immediately resulted in a perfectly predictable effect. However, considering our daily experiences, it quickly becomes clear that this is different for actions in everyday life. It might be that the same action results in the desired effect on the one day, while it results in another, unexpected effect, or no effect at all on the other day. Furthermore, not every effect occurs immediately after the action was carried out but might come with some temporal delay. Imagine, for example, the following situation: The room you have just entered is dark, so you decide to press the light switch in order to enlighten the room. What you usually would expect is that the light source would turn on. However, sometimes, the room might remain dark because the light bulb is broken or a thunderstorm caused a power blackout. If you use energy saving bulbs, the room will eventually get bright, but only some seconds after pressing the light switch. Given these observations, the question arises of whether and how effect-based action-effect learning is affected by such deviations from a “perfect” (i.e., immediate and perfectly contingent) relationship between actions and their effects. In the following two paragraphs, I will provide a short overview on the impact of effect contingency and effect contingency, two factors that are known to moderate learning mechanisms, and their relation to the acquisition of action-effect associations.

Human behavior is characterized by the ability to perform goal-oriented actions, which eventually allows humans to change their environment in a way they had in mind

---

before executing the action (Prinz, 1997). For example, we unlock our smartphone in order to make a phone call, or press the button on the coffee machine to get a cup of coffee. However, this ability requires a deep understanding of the complex connections within the world, and it is not innate, but rather the result of exploration and a continuous learning process which starts as early as from birth on (Piaget, 1952; Rochat, 1998). The associative knowledge about the consequences that follow an action under specific circumstances also affects adult action control as demonstrated by findings on the acquisition of action-effect associations (Beckers, Houwer, & Eelen, 2002; Elsner & Hommel, 2001; Hoffmann et al., 2009; Hommel, 1996; Hommel et al., 2003; Pfister et al., 2011; Wolfensteller & Ruge, 2011). This knowledge can be considered as a prerequisite to reach a certain goal in a specific situation. In traditional associative learning accounts, an association is defined as a mental link which connects two events (Pearce & Hall, 1980; Rescorla, 1968, 1988; Rescorla & Wagner, 1972). Furthermore, it is known that the strength of this link is moderated by the constant co-occurrence of two events – that is, *contingency* – and the temporal and spatial proximity of two events – that is, *contiguity*. If one considers associative learning mechanisms as the basis for action-effect learning, one could suggest that the acquisition of action-effect associations is sensitive to the same moderators that have been shown to exert an influence on associative learning mechanisms.

This was exactly the starting point for the seminal study of Elsner and Hommel (2004) on learning mechanisms in ideomotor actions, which provided evidence that action-effect learning is subject to the same limits as demonstrated for associative learning mechanisms in general. The authors relied on an adapted version of the two-phase action-effect acquisition paradigm (see *Section 2.1.2.1* for details with respect to the paradigm). In brief, they observed that actions in terms of key presses became associated with their effect tones only when contingency was high, that is, when performing an action predictably increased the probability of the associated effect tone to occur. The authors further demonstrated that

the acquisition of action-effect associations was dependent on the temporal contiguity between action and effect, such that action-effect associations were only formed when the temporal delay between an action and its effect did not exceed a certain threshold: Action-effect learning was evident at short delays (50 ms), but not when the delay was rather long (2s). The authors proposed a time window of up to one second between action and effect as the critical period for action-effect learning. If an effect occurred out of range of this critical time window, it would not be associated with an action. Thus, temporal proximity of action and effect has shown to foster action-effect learning when the time window is relatively small (Elsner & Hommel, 2004). However, the results of Dignath et al. (2014) challenge the view that effective action-effect learning is restricted to a time window as short as one second, as they observed evidence for the acquisition of action-effect associations in a series of high-powered experiments with temporal effect delays of up to two seconds.

Moreover, the temporal interval between action and effect is not only a moderator variable that promotes or inhibits the acquisition of action-effect associations, but might also be important and informative in itself. Thus, it seems reasonable to assume that this temporal information can also be learned. First hints that temporal information also becomes an integral part of the action representation and thereby affects response generation were collected by Kunde (2003) and Kiesel and Hoffmann (2004). Briefly, the authors of both studies observed that effect duration influenced RTs, since response initiation was slowed down when actions were followed by long as compared to short effects (e.g., 240 ms vs. 80 ms in Kunde, 2003). Furthermore, the study of Haering and Kiesel (2012) demonstrated that participants learn and exploit temporal action-effect regularities for action control. Their results showed that participants learned action-specific delays, because participants responded faster to the onset of an action effect whenever it occurred at its associated delay rather than earlier or later than expected. The seminal study of Dignath et al. (2014, Experiment 3) provided direct evidence for the idea that the temporal delay between an action and its effect becomes part

---

of the acquired action-effect associations in a bi-directional manner. In this study, participants experienced in an acquisition phase that actions consistently produced either immediate or delayed effect tones (50 ms vs. 2,000 ms action-effect delay). In a subsequent test phase, the former effect tones were presented as imperative stimuli. Their results showed that RTs in response to the imperative stimuli were reduced for short relative to long effect delays, which suggests that the information about the temporal delay between an action and its effect is not only learned, but also retrieved during action control. In line with ideomotor theory, they proposed that humans acquire bi-directional action-delay-effect associations. When planning an action, the anticipation of the intended effect should automatically reactivate the associated temporal delay as well as the corresponding action, and thereby influence response initiation. The finding that the action-effect delay is retrieved for action control has been corroborated and specified by Dignath and Janczyk (2017). Their results suggest that humans indeed learn *when* an action effect occurs, but that this temporal information can be acquired and used for action control without any knowledge about the identity of the effect.

## 2.3 EFFECT-BASED CONTROL OF EYE MOVEMENTS

### 2.3.1 IDEOMOTOR MECHANISMS IN GAZE CONTROL

For a long time, eye movements were largely neglected as an action domain in ideomotor control research. This is likely because we do not readily think of oculomotor actions (i.e., saccades; in the following oculomotor actions and saccades are used as synonyms) as typical means to generate effects in the environment to achieve a specific goal, even though recent technologies support that aspect, e.g., gaze-based software menu control, camera autofocus, or gaze-based communication software for motor impaired patients (Slobodenyuk,

---

2016). Several studies have shown that expected reward affects where people look (Bucker, Silvis, Donk, & Theeuwes, 2015; Hickey & van Zoest, 2012; Schütz, Trommershäuser, & Gegenfurtner, 2012; Theeuwes & Belopolsky, 2012) as well as how quickly they do so (Dunne, Ellison, & Smith, 2015; Lauwereyns, Watanabe, Coe, & Hikosaka, 2002; Milstein & Dorris, 2007; Rothkirch, Ostendorf, Sax, & Sterzer, 2013; Watanabe, Lauwereyns, & Hikosaka, 2003). Hence, reward can serve as a goal that determines oculomotor actions. Related to that, Sprague and Ballard (2003) have proposed a model of eye-movement control on how the visual system integrates reward maximization and task demands within a visuo-motor task with multiple competing goals. However, the authors considered such consequences of eye movements within their model in terms of retrieving task-relevant information from the environment, while the focus of the present work is on the control of eye movements that serve the goal of affecting the gaze behavior of other people (which, however, might of course also imply informational and thereby reward-laden value).

With respect to the ideomotor approach, some studies have shown that fundamental cognitive principles underlying other action control domains (e.g., manual, vocal) also apply to eye movements (Bompas, Hedge, & Sumner, 2017; Huestegge, 2011; Pieczykolan & Huestegge, 2014; Pieczykolan & Huestegge, 2017). By contrast, until a few years ago it has only rarely been addressed whether goal-based voluntary eye-movement control relies on the same basic cognitive mechanisms that drive goal-based action control in more standard effector systems (e.g., manual). In this context, the development of action control in infants using gaze-contingent paradigms was investigated. These studies addressed, for example, oculomotor reinforcement learning (Vernetti, Smith, & Senju, 2017) and mechanisms of controlling the environment via gaze (Verschoor, Paulus, Spape, Biro, & Hommel, 2015; Wang et al., 2012; Wass, Porayska-Pomsta, & Johnson, 2011). These studies generally show that infants are already able to anticipate oculomotor action outcomes and to control their environment using their eyes. A crucial advantage of focusing on gaze when testing very young



---

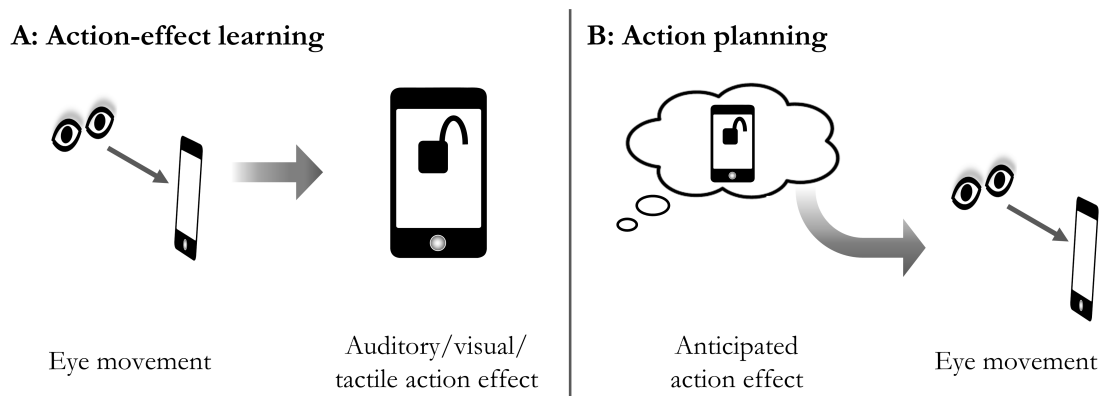
infants is that infants as young as four months are already capable of visuo-attentional control (Johnson, Posner, & Rothbart, 1991), while motor control is still limited at that age. Infants as young as seven months are able to acquire oculomotor action-effect associations, while the use of such action-effect associations for action control has been shown to occur at an age of around one year (Verschoor, Spape, Biro, & Hommel, 2013).<sup>3</sup>

However, more relevant for the present work are studies focusing on effect-based oculomotor control in adults. These studies have extended the ideomotor approach to the oculomotor domain using saccade latencies, the oculomotor counterpart to manual RTs, to measure action-effect learning (e.g., Herwig & Horstmann, 2011; Huestegge & Kreutzfeldt, 2012; Verschoor et al., 2013; see Herwig, 2015, for a recent review). Among them, the study by Huestegge and Kreutzfeldt (2012) is of particular relevance for the presented research. The authors observed that models of eye-movement control (e.g., Findlay & Walker, 1999), unlike motor control models in other action domains, tend to neglect the role of anticipated saccadic action effects (c.f. ideomotor principle, see *Section 2.1.1*) in gaze control. Huestegge and Kreutzfeldt (2012) then tested whether learned associations between saccades (toward targets in the left or right hemifield) and their arbitrary effects in the inanimate environment (gaze-contingent appearance of a square or diamond, respectively) affected saccade control. Results showed that presenting a task-irrelevant stimulus, which was either congruent, incongruent, or unrelated to the subsequent action effect, prior to an imperative auditory stimulus influenced saccade performance. Specifically, saccade latencies were shorter and accuracy was higher in congruent (vs. incongruent) conditions, indicating that internal representations of (oculomotor) action effects are generated prior to action execution and affect gaze

---

<sup>3</sup> Note that the spontaneous acquisition and bi-directional use of action-effect associations has also been demonstrated for manual actions in infants as young as nine months (Verschoor, Weidema, Biro, & Hommel, 2010). Relatedly, Eenshuistra, Weidema, and Hommel (2004) adapted the action-effect acquisition paradigm (Elsner & Hommel, 2001) such that children from the age of four years could be tested. Together, these studies demonstrate that the core assumptions of the ideomotor theory apply not only to adults, but also to children and infants, with infants showing earlier signs of effect-based action control for eye movements than for manual actions.

control. Overall, these studies showed that people acquire bi-directional associations between oculomotor actions and their visual effects, which are then retrieved for saccade control (see *Figure 2*). This idea, namely that anticipations shape oculomotor behavior, is further supported by recent findings indicating that acquired non-social effects of oculomotor actions affect visual search performance (Herwig & Schneider, 2014; Weiss, Schneider, & Herwig, 2014).



*Figure 2.* A model of ideomotor gaze control as an extension of the ideomotor theory of action control to the oculomotor domain.

### 2.3.2 SOCIOMOTOR MECHANISMS IN GAZE CONTROL

When investigating goal-based behavioral control of eye movements from an ideomotor perspective, however, one has to consider gaze as an action modality, which has direct consequences not only in the inanimate, but also in the animate environment, that is in other persons. To give a similar everyday example as the pointing example in *Section 2.1.3*, person A can use her/his gaze to guide the attention of a person B toward an interesting object or location by directing the gaze toward that object or location (gaze cueing; see Frischen, Bayliss, & Tipper, 2007). Observing such an eye movement will almost automatically prompt person B to look at the object or location as indicated by the gaze direction of person A. Interestingly, Kunde et al. (2018) even go a step further by claiming that gaze could be regarded as a specialized effector system in social contexts, which has evolved to ensure efficient social action control. They argue that, besides their main task of visual perception, the

---

eyes are more or less reserved to affect other's behavior, while they are less useful to manipulate the inanimate world, assuming a special relationship between the oculomotor system and social responses. Thus, social interaction represents a prime example for environmental control by eye movements. According to the idea of Kunde et al. (2018), this context-specific applicability distinguishes eye movements from other effector systems, such as arm movements. In contrast to eye movements, arm movements can serve to affect both the inanimate as well as the animate environment, when arms are used either to move an object from A to B or to communicate something to a friend using a gesture. Before I will further describe anticipation-based control of eye movements in a social context, I am going to summarize important findings on the role of gaze in social contexts.

The eyes have been shown to be an important tool for exchanging – receiving as well as providing – information with other humans in many different situations (Gobel, Kim, & Richardson, 2015), and there has been extensive research on the role of gaze in social contexts (Frischen et al., 2007). It has been demonstrated, for example, that direct eye contact is highly attention-grabbing (Böckler, van der Wel, & Welsh, 2014, 2015; Boyer & Wang, 2018; van der Wel, Welsh, & Böckler, 2018), and modulates cognitive processes related to face processing (e.g., Grünau & Anston, 1995; Hood, Macrae, Cole-Davies, & Dias, 2003; Macrae, Hood, Milne, Rowe, & Mason, 2002; Senju, Hasegawa, & Tojo, 2005), judgements of attractiveness (Bayliss, Paul, Cannon, & Tipper, 2006; Jones, DeBruine, Little, Conway, & Feinberg, 2006), deception (e.g., Mann et al., 2012), or prosocial behavior (e.g., Haley & Fessler, 2005). Perceiving averted gaze has been shown to shape gaze behavior in a powerful manner as it automatically leads visual attention toward the perceived gaze direction (Driver et al., 1999; Friesen & Kingstone, 1998; Frischen et al., 2007; Mansfield, Farroni, & Johnson, 2003). In the study by Mansfield et al. (2003), participants observing averted gaze were unable to suppress shifting their gaze into the perceived gaze direction even though they were instructed to do so. The automatic and reflexive nature of gaze cueing suggests that observing

---

another's gaze shift evokes a similar (simulated) motor response in the observer, as captured by the mirror-neuron system (Grafton, Arbib, Fadiga, & Rizzolatti, 1996; di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; see Rizzolatti & Craighero, 2004, for a review). In line with this view, similar brain regions have shown to be involved in both executing and observing eye movements (Grosbras, Laird, & Paus, 2005). Amazingly enough, the above-mentioned sensitivity to eye gaze has shown to be of an innate nature such that already newborns are able to discriminate direct from averted gaze (Farroni, Csibra, Simion, & Johnson, 2002; Farroni, Massaccesi, Pividori, & Johnson, 2004). It has also been demonstrated that infants as young as three to four months (or younger: see Farroni, Massaccesi et al., 2004) use gaze cues to direct their attention (Farroni, Johnson, Brockbank, & Simion, 2000; Farroni, Johnson, & Csibra, 2004). Another field of research addressed effects of observed gaze direction on the perception of emotional face expressions. For example, the activity of the amygdala in response to faces displaying anger or fear was modulated by gaze direction (Adams, Gordon, Baird, Ambady, & Kleck, 2003). Amygdala activity was especially pronounced for angry faces with averted gaze or fearful faces with direct gaze (vs. vice versa). To explain these results, it has been suggested that for the emotional expressions of avoidance (e.g., fear, sadness) versus approach (e.g., anger, joy), perception is enhanced by averted versus direct gaze (Adams & Kleck, 2005).

However, these studies predominantly used gaze as a social stimulus to trigger behavioral responses in others. While still focusing on the gaze recipient, the study of Joyce, Schenke, Bayliss, and Bach (2016) pursued a slightly different approach to the investigation of gaze by asking whether gaze cueing is predictive. More specifically, participants had to categorize objects appearing to the left and right side of the screen as belonging to categories such as foods or drinks, while a centrally presented face either looked at the object-to-be-classified, at the other object, or straight ahead. The crucial manipulation was that one face (e.g., the male face) looked consistently at items of one category (e.g., food) and the other

---

face (e.g., the female face) consistently looked at items of the other category (e.g., drinks). The results suggested that gaze cueing is indeed predictive as participants picked up these regularities, and that the corresponding expectations elicited shifts of attention resulting in faster object categorization for objects the particular face usually looks at when the faces looked straight ahead. These results implicate that expectations about the other's gaze have the power to shape attention processes. While this finding is informative with respect to anticipatory gaze cueing, it still does not address the cognitive mechanisms underlying anticipation-based gaze interaction from the viewpoint of the actor, that is, how eye movements that elicit behavioral changes in other people, are brought about.

Given that the processing of social information is assumed to differ from the processing of other kinds of information and given the special role of gaze in social contexts, as outlined above, it is reasonable to ask whether effect-based control of eye movements is qualitatively different for social versus non-social action contexts, or whether similar underlying mechanisms guide gaze control in both social and non-social contexts (see Kunde et al., 2018). Similar deliberations were also a starting point of the study by Herwig and Horstmann (2011) who were the first to demonstrate acquisition as well as anticipation of social action effects for the oculomotor system. Participants first learned that a saccade toward either of two emotionally neutral faces on the left or right of a central fixation cross triggered the onset of an emotional expression (e.g., left face: happy/right face: angry) in these faces (i.e., the effect) with a temporal delay of 100 ms. In a later test phase, participants were presented with the former effect stimuli (happy/angry faces) in the center with the instruction to saccade to the left or right. Crucially, one group had to respond with a saccade to the position associated with the particular facial expression in the learning phase (i.e., action-effect congruent mapping), while instructions were reversed for the other group (i.e., action-effect incongruent mapping). Faster saccade latencies for the group with the action-effect congruent mapping indicated that participants indeed acquired associations between their

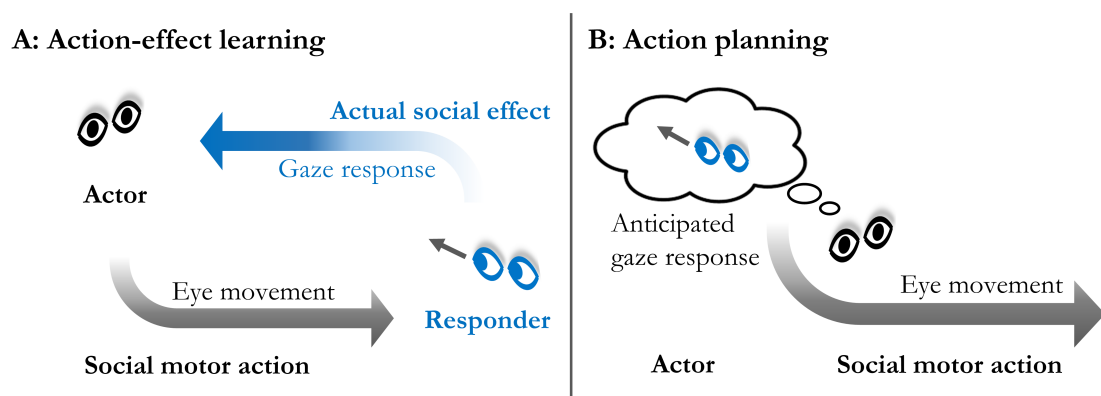
eye movement and the triggered change of facial expression. Furthermore, an additional analysis of saccade landing positions during the learning phase showed that participants tended to direct their initial saccade toward the location in the face where the most dominant changes occurred, that is, to the mouth region for happy faces, and to the eyebrow region for angry faces. This difference in the landing position between the two effect conditions (angry vs. happy face) increased over the course of the experimental blocks (Herwig & Horstmann, 2011). The authors interpreted this finding as evidence for the assumption that saccades are selected by the anticipation of their effects in a social context.

In summary, one can conclude that there is a large number of studies on gaze-related topics, for example on gaze processing, gaze cueing, or the development of eye gaze, while the number of studies investigating gaze control from the viewpoint of the actor is rather negligible. There are, for example, several studies from the Bayliss lab (e.g., Bayliss et al., 2013; Edwards, Stephenson, Dalmaso, & Bayliss, 2015; Stephenson, Edwards, Howard, & Bayliss, 2018; see *Section 6.1* for a more detailed discussion of these studies) which focus on the gaze initiator in a gaze-interaction scenario. While these studies are very informative with respect to behavioral and evaluative consequences of having one's gaze followed or not in a gaze-interaction scenario, they are not intended to explain the cognitive foundations of initiating eye movements in such a scenario – a research question which I am focusing on in the course of this work.

## 2.4 OVERVIEW OF THE PRESENT WORK

As outlined in the previous sections, recent research has investigated ideomotor control processes in eye movements, demonstrating that the acquisition (Herwig & Horstmann, 2011, for the social domain; Huestegge & Kreutzfeldt, 2012, for the non-social domain) and anticipation (Herwig & Horstmann, 2011, for the social domain) of associations

between between oculomotor actions and their visual effects play an important role in saccade control. These results suggest that when planning a goal-oriented eye movement, the anticipation of the intended effect (i.e., the goal) activates the saccade that produced the intended effect during previous experiences. One of the most apparent situations where eye movements are deliberately used to influence another person's behavior is when one person performs an eye movement that is intended to elicit a certain gaze response in the other person, for example toward a potential source of threat or reward. In a bar, for example, an eye movement can serve to draw the attention of your conversation partner to the seemingly dubious person in the back corner of the bar, or to the attractive person sitting at the next table. Given the relevance of such gaze-interaction scenarios in everyday life, it is reasonable to focus on saccade control in a gaze-interaction situation from the viewpoint of sociomotor gaze control. Based on the results of Herwig and Horstmann (2011) and the deliberations of Kunde et al. (2018), I propose that such eye movements are also represented in an effect-based manner (see *Figure 3*), thereby representing an integral part of ideomotor gaze control. Therefore, the following series of experiments of this work was particularly designed to systematically address cognitive foundations of intentional eye-movement control in a gaze-interaction scenario.



*Figure 3.* Proposed model of a sociomotor loop for eye movements: Actors learn which eye movement evokes what kind of gaze response in the responder, thereby acquiring bi-directional action-effect associations (A). After learning, the anticipation of the intended gaze response reactivates the corresponding eye movement (B). Figure adapted from Kunde et al. (2018).

After the general introduction into the broader scientific field of ideomotor theory of action control, the following chapters will focus on the empirical part of this work. In the first empirical part (Chapter 3), I describe two approaches to the issue of effect-based oculomotor control both of which are designated to identify similarities and differences in mechanisms underlying ideomotor gaze control in a social versus non-social context. The first approach (Experiment 1) focuses on the anticipation of upcoming action effects in action control by asking whether action-effect anticipation is reflected in corresponding spatial saccade characteristics in inanimate environments as reported for animate environments (cf. Herwig & Horstmann, 2011). The second approach (Experiment 2–6) targets the acquisition of bi-directional action-effect associations in a gaze-interaction scenario in direct comparison with equivalent, but non-social scenarios.

As described in *Section 2.2*, action-effect learning has an associative basis (Elsner & Hommel, 2004), and is therefore subject to all influences known to impact on associative learning, for example contingency and contiguity. If one adopts the general view of Kunde et al. (2018) that similar mechanisms of action control can be transferred to the social domain, sociomotor action-effect learning should also be influenced by the factors known to foster or hinder action-effect learning, like contingency and contiguity. However, since responses of an interaction partner are often variable and take some time to occur, humans may take these variable and delayed responses of an interaction partner into account, suggesting that the action-effect learning mechanisms can cope with these features differently in the social domain. The question of whether and how such naturally occurring deviations from a “perfect world” (i.e., a world perfectly contingent on our behavior and intentions) are incorporated into sociomotor action control is addressed in experiments described in Chapter 4 (Experiments 7 and 8). The next empirical part (Chapter 5, Experiments 9 and 10) compares two ways in which an action effect (i.e., the goal of the eye movement) is specified, namely in an intention-based mode, where participants are free to decide which action (of a



---

set of actions) to perform next (free-choice acquisition), or in a stimulus-based mode, where an external stimulus indicates the action to be performed next (forced-choice acquisition). These experiments compare the two modes of acquiring action-effect associations, that is, a free- or forced-choice mode, in ideomotor gaze control in a social versus non-social context. Together, Chapter 4 and 5 answer the question as to whether oculomotor action-effect learning is subject to the same processes and constraints in the acquisition of action-effect associations as demonstrated for manual action-effect learning. In the final empirical part (Chapter 6, Experiments 11 and 12), I establish a new gaze-interaction scenario closer to real-life interaction, which is especially suited to study the anticipation of action-effect associations under conditions of higher ecological validity.

Another essential research question I am addressing in this work is whether social environments are special in that they recruit different processing mechanisms and/or dynamics with respect to effect-based processing, that is, whether ideomotor gaze control differs for a social versus non-social context. According to Kunde et al. (2018), there are good reasons to hypothesize that mechanisms of gaze control are very similar for both contexts. The assumption of similar processes for social and non-social contexts is also reflected in recent findings on reflexive orienting. While early studies suggested that eyes may be an evolutionary important and therefore special cue to direct attention (Friesen & Kingstone, 1998), further research has stressed the similarities between gaze cueing and other, non-social ways of reflexive orienting, for example, based on peripheral cues or centrally presented arrows (see Frischen et al., 2007; Kuhn & Kingstone, 2009). However, social (vs. non-social) action effects might come with some peculiarities: For instance, social action effects might be less predictable and less immediate than non-social action effects (see *Section 4.1*, where the role of effect contingency and contiguity in social vs. non-social environments will be discussed in more detail). While there is evidence for effect-based control of eye movements from both an ideomotor and sociomotor point of view (Herwig & Horstmann, 2011; Huestegge

& Kreutzfeldt, 2012), a systematic test of whether similar processes can be assumed for both social and non-social contexts is still pending, which represents an essential research question for this work. This issue will be discussed in detail in the General discussion (Chapter 7).

---

## 3 IDEOMOTOR GAZE CONTROL IN SOCIAL VERSUS NON-SOCIAL CONTEXTS

While there is already a number of relevant studies on ideomotor principles in gaze control in adults (e.g., Herwig & Horstmann, 2011; Herwig & Schneider, 2014; Huestegge & Kreutzfeldt, 2012; Verschoor et al., 2013), there are still some open questions that have not yet been answered empirically, especially with respect to gaze control within animate versus inanimate environments. More specifically, a clear research gap exists with respect to the control of goal-oriented eye movements that serve to elicit changes in the gaze behavior of the interaction partner, such as when humans move their eyes in order to get the other person to look at a particular place. As outlined in *Section 2.4*, there are reasons to assume that such control processes in eye movements might be different due to peculiarities inherent to the social (vs. non-social) nature of the situation. Thus, the overall goal of the first empirical part of this dissertation is to address oculomotor gaze control in animate and inanimate environments. While the first experiment asks whether similar spatial saccade characteristics as revealed for animate environments (Herwig & Horstmann, 2011) can be observed for inanimate environments (Experiment 1) by exclusively considering a non-social context, the following series of experiments (Experiments 2–6) includes a systematic comparison between ideomotor and sociomotor gaze control within the identical paradigm.

## 3.1 EXPERIMENT 1

It is known that oculomotor action-effect associations, that is, links between saccades and the visual change that is triggered by such eye movements, are acquired and used for saccade control in a social context (Herwig & Horstmann, 2011). The use of such saccade-effect associations for oculomotor control was reflected in spatial saccade characteristics such that participants automatically directed their gaze toward the anticipated location of change. More specifically, participants directed their saccades more often toward the mouth region of a neutral face when the neutral face was about to change into a smile and toward the eyebrow region when the neutral face was about to change into an angry face. Given the many findings about action-effect associations in the non-social domain with manual actions (e.g., Elsner & Hommel, 2001; Herwig et al., 2007) the question arises whether similar results as observed in Herwig and Horstmann (2011) would also emerge with arbitrary action effects, such as changes in shape or color of non-social objects rather than changes in facial expression. While oculomotor action-effect associations have also been shown to play an important role in inanimate environments (Huestegge & Kreutzfeldt, 2012), the design of the Huestegge and Kreutzfeldt (2012) study did not allow for an analysis of spatial saccade characteristics as an index of anticipation. Thus, the question of whether saccade landing positions as a spatial measure of effect anticipation can reflect anticipation in an inanimate environment remains an open issue.

In the context of manual ideomotor control, research has already focused on dynamic aspects underlying the acquisition of action-effect associations, that is, regarding the temporal contiguity between actions and effects (*see Section 2.2*). Specifically, it was shown that the length of the temporal action-effect interval affected the latency of the (manual) action (Dignath et al., 2014). They explained this effect by assuming that temporal information about the action-effect interval is integrated into a cognitive action script (or event

file) in a bi-directional manner, which is then automatically retrieved during action preparation and thereby affects the timing of action selection. If this general interpretation is correct, one would expect a corresponding phenomenon in the oculomotor domain, a prediction that has not been tested yet. Furthermore, it is not known whether and how different temporal action-effect intervals affect the dynamic buildup of action-effect associations in oculomotor control. While the action effect was presented throughout the trial or rather occurred during saccade execution in the Huestegge and Kreutzfeldt (2012) study, the action effect occurred 100 ms after the saccade arrived at one of the faces in the study of Herwig and Horstmann (2011). Since eye movements differ from other action modalities (e.g., manual actions) in that they consist of fast sequences of jerky saccades and stationary fixations (Liversedge, Gilchrist, & Everling, 2011), it appears especially interesting to investigate this phenomenon in the oculomotor domain.

Thus, the central aspect of the present study was to narrow down basic mechanisms underlying the acquisition of action-effect learning in the oculomotor domain with respect to both spatial (by focusing on effects on spatial oculomotor control) and temporal aspects (by studying effects of temporal action-effect interval and by examining learning dynamics). To do so, I analyzed saccadic eye movements that contingently resulted in specific (immediate or delayed) changes in inanimate objects (traffic lights). The experimental paradigm (see *Section 3.1.1* for a detailed description of the setup) was a free-choice design that did not include a congruency manipulation, but rather focused on evidence of spatial effect location anticipation in eye movements. The pictures of two neutral identical traffic lights (i.e., lights without any lamp turned on) were presented on a screen, one on the left and one on the right side of central fixation. Contingent upon the participant's freely selected gaze direction (toward the left versus right traffic light), one lamp of the light turned on (e.g., upper lamp when the left traffic light was targeted vs. lower lamp when the right traffic light was targeted). The

temporal onset of the action effect was either immediate (0 ms action-effect interval) or delayed (100 ms or 300 ms).

Experiment 1 addressed three novel research questions. First and foremost, I tested whether action-effect anticipation in oculomotor control is reflected in corresponding spatial saccade characteristics in inanimate environments (similar to corresponding effects demonstrated in a social context, see Herwig & Horstmann, 2011). If action-effect learning occurs and thus saccade preparation toward the peripheral target involves the mental representation of the action effect, participants should anticipate the specific change in the neutral target objects and direct the landing position of their initial saccade toward the anticipated effect location (upper/lower lamp of traffic light). Second, I tested whether the previously reported dependency of action latency on the temporal effect delay (action-effect interval, see Dignath et al., 2014) also occurs in the oculomotor domain. Third, I asked which temporal effect delay is optimally suited to develop strong associations between oculomotor actions and their visual effects over the course of the experiment. While effects of our actions in the inanimate environment usually occur instantaneously (e.g., in the case of light switches), thus supporting the prediction that immediate effects should be optimal for learning, it is also possible that immediate visual changes in the oculomotor domain (i.e., changes occurring during the saccade toward the target) are less salient than delayed changes due to saccadic suppression phenomena, which would rather support the prediction of enhanced learning with delayed effects.<sup>4</sup>

---

<sup>4</sup> Note that my individual contribution to Experiment 1 was the analysis and interpretation of data that had already been collected, as well as manuscript writing (see *Appendix B* for a detailed statement of individual author contributions).

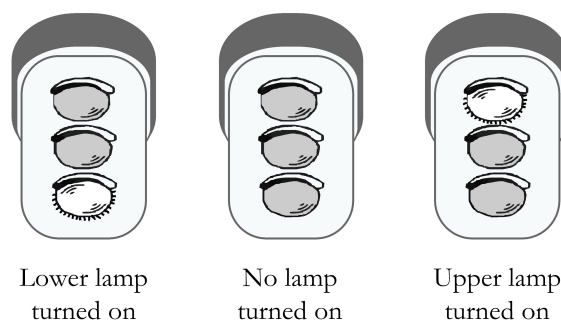
### 3.1.1 METHODS

#### 3.1.1.1 Participants

Forty-eight students took part in the experiment (mean age = 24.2 years, standard deviation ( $SD$ ) = 4.3 years, age range: 18 – 41 years, 13 male) and received course credits. All participants reported normal or corrected-to-normal vision and were naïve with respect to the purpose of the experiment. They gave their informed consent before the experiment was started.

#### 3.1.1.2 Stimuli

The stimulus type used in the present study (i.e., traffic lights) is depicted in *Figure 4*. There were three different versions of the traffic light stimulus (no lamp vs. upper lamp vs. lower lamp turned on) that differed only in the upper or the lower lamp region (i.e., the upper or lower light was switched on or off). The (greyscale) stimuli were presented on a white background. The size of each stimulus was  $5.09^\circ \times 7.36^\circ$  (width  $\times$  height) of visual angle. Stimuli were located to the left and right of a black fixation cross at a horizontal distance of  $8^\circ$  visual angle. The size of the fixation cross amounted to  $0.4^\circ \times 0.4^\circ$  of visual angle ( $14 \times 14$  pixels).



*Figure 4.* Stimulus material used in the current experiment, showing schematic (greyscale) representations of traffic lights: a light with the upper lamp turned on (left), a neutral light without any lamp turned on (middle), and a light with the lower lamp turned on (right). Figure adapted from Ref. [1].

### 3.1.1.3 Apparatus, task and procedure

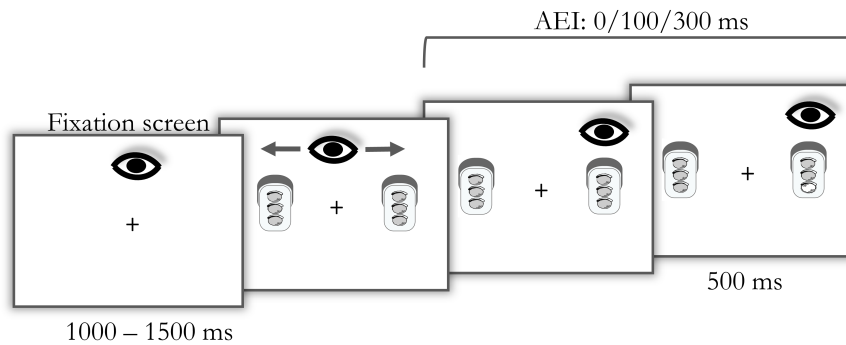
The experiment was performed in a dimly lit room with a viewing distance of 71 cm (Bielefeld) or 62 cm (Aachen) to the display monitor (refresh rate: 100 Hz). Screen resolution was set to  $1024 \times 768$  pixels on a  $36 \text{ cm} \times 27 \text{ cm}$  (width  $\times$  height) screen (Bielefeld) or  $41 \text{ cm} \times 30 \text{ cm}$  (Aachen). Eye movements of the right eye were recorded with a video-based eye tracker (Eye Link, SR Research, Ontario, Canada). Half of the participants were recorded using an EyeLink 1000 (1000 Hz sampling rate; Bielefeld), while the other half were recorded using an Eye Link II system (at 500 Hz sampling rate; Aachen). Head and chin were stabilized by a forehead and chin rest, respectively. The experiment was programmed using Experiment Builder (SR Research, Ontario, Canada).

Participants received a visual instruction prior to the onset of the four experimental blocks involving 56 trials each (224 trials in total). A calibration of the eye tracker was performed prior to each experimental block. Each trial started with a central fixation cross (see *Figure 5*). The length of the fixation interval varied randomly between 1000–1500 ms. Then, two neutral traffic light stimuli appeared simultaneously to the left and right of the fixation cross. Participants were instructed to look at the fixation cross and then to freely decide for one of the two (left/right) objects as a saccade target as soon as these objects appeared. Participants were told to avoid possible preferences in gaze or regularities in saccade sequences. Every 56 trials, feedback was provided regarding the number of location choices (left vs. right) to ensure that the experience of each saccade-effect combination was equally balanced. Each saccade triggered a particular change (turning-on of the upper/lower lamp) within the fixated neutral traffic light. The time of the onset of the stimulus change after the arrival of the initial saccade at the effect location, that is, the action-effect interval (AEI), was manipulated between participants in three stages (0 ms vs. 100 ms vs. 300 ms). Note that for the 0 ms AEI condition, stimulus change after target presentation was initiated by the first saccade leaving the interest area around the fixation cross to ensure a change



during saccade execution. The change lasted for 500 ms (for the 0 ms group after saccade landing), after which participants were instructed to reorient their gaze toward the screen center.

Crucially, the stimulus' change depended on the saccade's direction: For half of the participants, a leftward saccade triggered a change from a neutral light to a light with the upper lamp turned on, and a rightward saccade triggered a change from a neutral light to a light with the lower lamp turned on. For the other half of participants, this saccade-effect mapping was reversed. No explicit information about the mapping was provided to the participants.



*Figure 5.* Schematic representation of a trial: After the presentation of a white screen with a black fixation cross (randomized duration of 1000–1500 ms), two identical light stimuli appeared simultaneously to the left and right of the fixation cross. The participant's freely chosen saccade to one of the stimuli (left/right) triggered a change within the fixated light stimulus which lasted for 500 ms. Per each action-effect interval (AEI) condition, a total of 224 trials were presented. Figure adapted from Ref. [1].

#### 3.1.1.4 Design

To analyze the data, I applied repeated-measures analyses of variance (ANOVA,  $\alpha = .05$ , throughout) together with  $\eta^2_p$  as effect size estimates. Sphericity violations were assessed using the Mauchly's sphericity test. In case of sphericity violations, I reported Greenhouse-Geisser corrected  $p$ -values along with original degrees of freedoms as well as corresponding  $\epsilon$ -values. I conducted additional ANOVAs with a reduced number of factors to

break down relevant significant interaction effects. Two-tailed paired *t*-tests were applied for follow-up comparisons between conditions.<sup>5</sup>

The present experiment utilized a mixed design with *block* (four separated blocks of equal length) as a within-subjects variable and *AEI* (0 ms vs. 100 ms vs. 300 ms) as a between-subjects variable. Spatial and temporal saccade parameters were recorded. For spatial analyses, I computed relative vertical landing positions (rVLP), that is, the difference (measured in degrees of visual angle) between the mean upper (in upper lamp condition) and the mean lower (in lower lamp condition) vertical saccade landing position, indicating a spatial deviation of the saccade landing position toward the effect location. Thus, positive rVLP values were indicators for spatial effect anticipation (irrespective of actual lamp position). I additionally analyzed whether already the initial saccade toward the peripheral target reflected maximal anticipation of the effect location, or whether the saccade after the initial saccade is directed even closer to the effect location. The latter would indicate that anticipation is reflected in a sequence of saccades rather than being restricted to the initial saccade. To address this issue, I calculated the distance (measured in degrees of visual angle) between the mean vertical landing position of the saccade (both the initial saccade and the subsequent saccade) and the vertical position of the enlightened effect location (analyzed separately for the upper and lower lamp condition). If a smaller deviation for the second saccade (compared with the initial saccade) was found, this would indicate that anticipation is reflected in a sequence of saccades. For the temporal analyses, I analyzed saccade latency, which represents the oculomotor counterpart to manual RTs and was defined as the interval between the onset of the light stimuli and the initiation of the saccade to one of the two targets.

---

<sup>5</sup> Note that this procedure applies to any experiment reported in the present work.

### 3.1.1.5 Additional remark

Note that the actual study included four additional blocks of 56 trials each involving schematic face stimuli (as an alternative to the inanimate traffic lights) that changed emotional expression similar to the photographic face stimuli used in Herwig and Horstmann (2011). The sequence of blocks involving face stimuli was presented either before or after the sequence of blocks involving traffic light stimuli (counterbalanced across participants). However, no empirical evidence of effect anticipation was found for these newly designed schematic stimuli, most likely due to the fact that the changes were less salient than in the photographs used in Herwig and Horstmann (2011), where anticipation effects were shown in natural, photographic face stimuli. The failure to find effect anticipation in these schematic face stimuli might also have its root in their unnatural character. To ensure comparability, the face stimuli were designed comparable to the traffic light stimuli with respect to the limited location of the change (i.e., the effect): The change of emotional expression in the face was restricted to either the eyebrow (c.f. the upper lamp change) or the mouth (c.f. the lower lamp change) region. Since it is well known that emotional expressions are not limited to a specific face region, the design decision may have induced a somewhat strange, unnatural appearance for these stimuli. Since the observation of spatial effect anticipation in oculomotor control is a prerequisite for addressing the specific present research questions, I here only focus on reporting the results for the traffic light stimuli. However, the role of stimulus type presentation order is addressed in the results section.<sup>6</sup>

---

<sup>6</sup> A depiction of the face stimulus material as well as a detailed analysis of the face stimulus data set is provided in the supplementary material of the previously published work: Reference [1]: Riechelmann, E., Pieczykolan, A., Horstmann, G., Herwig, A., & Huestegge, L. (2017). Spatio-temporal dynamics of action-effect associations in oculomotor control. *Acta Psychologica*, 180, 130–136. <https://doi.org/10.1016/j.actpsy.2017.09.003>

## 3.1.2 RESULTS

### 3.1.2.1 Distribution of choice frequencies

The distribution of choice frequencies of left- and rightward saccades was close to the instructed balanced distribution. The average proportion of left versus right saccades of all valid trials ( $N = 10181$ ) amounted to 49.93 % versus 50.07 %, respectively. No consistent preference in gaze or regularities in saccade sequences were apparent.

### 3.1.2.2 Spatial oculomotor performance

Trials in which the action effect was not triggered by a saccade within 1000 ms after onset of the light stimuli (3.18 %, equivalent to 342 of 10752 trials) were excluded from analysis. Furthermore, trials with more than one saccade within the critical time interval prior to effect presentation were excluded, as it was not possible to unambiguously determine the saccade which triggered the effect in these trials (0.13 %, equivalent to 14 of 10752 trials). In addition, only saccades with latencies longer than 99 ms to control for anticipatory saccades (2.07 % anticipatory saccades, equivalent to 215 of 10396 trials) were included, resulting in 10181 valid trials in total. The rVLPs were submitted to a two-way mixed ANOVA with block as a within-subjects factor and AEI as group factor (see *Figure 6*). There was a significant main effect of block on rVLP,  $F(3, 135) = 8.12$ ,  $\varepsilon = .71$ ,  $p < .001$ ,  $\eta^2_p = .15$ . The rVLP was smallest in the first block (mean ( $M$ ) =  $0.64^\circ$ , standard error ( $SE$ ) =  $0.12^\circ$ ), increased in the second and third block ( $M = 0.98^\circ$ ,  $SE = 0.14^\circ$ ;  $M = 1.07^\circ$ ,  $SE = 0.14^\circ$ , respectively) and remained nearly constant in the last block ( $M = 1.05^\circ$ ,  $SE = 0.14^\circ$ ). Pairwise comparisons revealed that this significant main effect of block mainly resulted from the significant difference between the first block when compared to the second,  $p = .001$ , the third,  $p < .001$ , and the fourth block,  $p = .002$ . The second compared to the third and fourth block, as well as the third and fourth block did not significantly differ (all  $ps > .05$ ).

Block significantly interacted with AEI,  $F(6, 135) = 3.11$ ,  $\varepsilon = .71$ ,  $p = .017$ ,  $\eta^2_p = .12$ . To further qualify this interaction, I conducted separate one-way ANOVAs for each block condition. While there was no significant effect of AEI in Blocks 1–3, all  $p$ s  $> .05$ , the AEI significantly affected the rVLP in Block 4,  $F(2, 45) = 3.40$ ,  $p = .042$ ,  $\eta^2_p = .13$ . In this block, the AEI = 0 ms condition showed a significantly smaller effect on rVLP compared to the AEI = 100 ms condition,  $p = .014$ . The AEI = 300 ms condition was not significantly different from both the AEI = 0 ms condition,  $p = .084$ , and the AEI = 100 ms condition,  $p = .440$ . For every experimental block, the mean value of the rVLP (averaged across AEI conditions) was significantly different from zero (all  $p$ s  $< .001$ ). The main effect of AEI was not significant,  $F(2, 45) = 1.62$ ,  $p = .210$ ,  $\eta^2_p = .07$ .

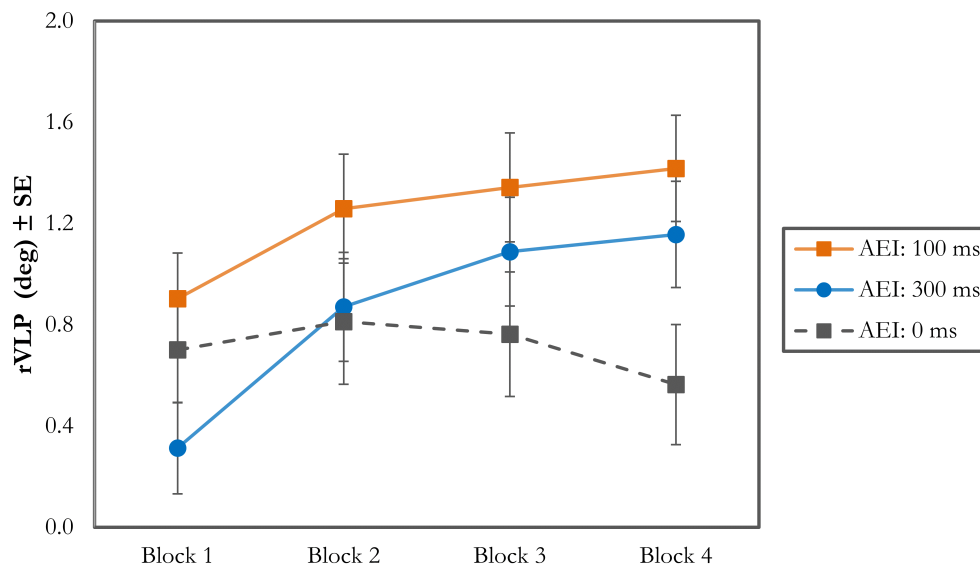


Figure 6. Mean relative vertical landing position (rVLP in  $^{\circ}$ ) as a function of action-effect interval (AEI: 0 ms vs. 100 ms vs. 300 ms) and experimental block (1–4). Error bars depict standard errors of the mean ( $SE$  in  $^{\circ}$ ). Figure adapted from Ref. [1].

To address the question of whether anticipation is reflected in the initial saccade or, alternatively, in a sequence of (two) consecutive saccades, two three-factorial repeated-measures ANOVAs (for upper/lower lamp condition) were conducted with block and saccade index (initial vs. subsequent saccade) as within-subjects factors and AEI as group factor. For the upper lamp condition, the main effect of saccade index was not significant,  $F < 1$ .

Also, there were no significant interactions with saccade index: There was neither a significant three-way interaction,  $F(6, 135) = 1.10, p = .365, \eta^2_p = .05$ , nor a significant two-way interaction of saccade index and block,  $F(3, 135) = 1.33, p = .266, \eta^2_p = .03$ , or of saccade index and AEI,  $F < 1$ . In the lower lamp condition, saccade index significantly affected rVLPs,  $F(1, 45) = 4.69, p = .036, \eta^2_p = .09$ . There was a greater deviation for the subsequent saccade ( $M = 1.39^\circ, SD = 0.12^\circ$ ) compared to the initial saccade ( $M = 1.14^\circ, SD = 0.11^\circ$ ). Further, the interaction of saccade index and block was significant,  $F(3, 135) = 11.74, p < .001, \eta^2_p = .21$ . Pairwise comparisons revealed that rVLPs were not significantly different for the first and second saccade in Block 1,  $p = .406$ . From Block 2,  $p = .060$ , to Block 3,  $p = .010$ , and to Block 4,  $p = .002$ , the difference between the initial and subsequent saccade increased, indicating smaller deviations from the effect location for the initial, but greater deviations from the effect location for the subsequent saccade. Thus, with their subsequent saccade participants directed their gaze rather away from the effect location (in the lower lamp condition) or kept the same distance to the effect location (in the upper lamp condition).

To explore effects of stimulus type presentation order (faces first vs. traffic lights first), rVLP was submitted to a three-factorial repeated-measures ANOVA with block as within-subjects factor and stimulus type presentation order as well as AEI as group factors. The significant interaction of stimulus type presentation order and block,  $F(3, 126) = 4.39, \epsilon = .74, p = .012, \eta^2_p = .10$ , indicated an influence of presentation order, with greater anticipation effects for light stimuli when they were presented first compared to presenting face stimuli first. Crucially, pairwise comparisons for each block revealed that stimulus presentation order was irrelevant for rVLP in Block 1,  $p = .969$ , and Block 2,  $p = .090$ . The difference between the rVLP was close to significance in Block 3,  $p = .053$ , and significant in Block 4 only,  $p = .039$ . Importantly, the three-way interaction was not significant,  $F(6, 126) = 2.06, \epsilon$

$= .74$ ,  $p = .086$ ,  $\eta^2_p = .09$ , indicating that the order of stimulus type presentation did not differentially influence the rVLP depending on AEI and block.

Taken together, these results indicate fast learning of spatial action-effect contingency in all AEI conditions (i.e., already in the first block). However, this contingency affected saccade control in terms of stronger spatial biases predominantly in the two delayed effect conditions. While there was no significant change of rVLP across blocks in the AEI = 0 ms condition,  $F < 1$ , I observed significant linear trends in the 100 ms condition and the 300 ms condition,  $F(1, 15) = 5.06$ ,  $p = .040$ ,  $\eta^2_p = .25$  and  $F(1, 15) = 9.36$ ,  $p = .008$ ,  $\eta^2_p = .39$ , respectively. The sequential saccade analysis revealed that after their initial saccade, participants did not direct their gaze more closely to the effect location, suggesting that anticipation is reflected in the initial saccade, not in a sequence of successive saccades. Stimulus presentation order was shown to be irrelevant in the first block of each sequence with traffic light stimuli.

### 3.1.2.3 Temporal oculomotor performance

Saccade latencies were submitted to a mixed ANOVA with block as within-subjects factor and AEI as group factor. Overall, the mean saccade latency amounted to 209 ms ( $SE = 5$  ms). There was a significant main effect of AEI,  $F(2, 45) = 11.59$ ,  $p < .001$ ,  $\eta^2_p = .34$  (see *Figure 7*). On average, saccade latency was longest for an AEI of 300 ms ( $M = 245$  ms,  $SE = 9$  ms), but much shorter in both the 100 ms condition ( $M = 185$  ms,  $SE = 9$  ms) and the 0 ms condition ( $M = 197$  ms,  $SE = 9$  ms). There was a significant interaction of AEI and block,  $F(6, 135) = 3.23$ ,  $\epsilon = .81$ ,  $p = .010$ ,  $\eta^2_p = .13$ . When analyzing the four experimental blocks separately with one-way ANOVAs, there were selective differences between AEI conditions in every block, all  $ps < .05$ . In Block 1, the AEI = 300 ms condition did not significantly differ from the AEI = 0 ms condition,  $p = .165$ . Regarding the AEI = 100 ms condition, the difference to the AEI = 0 ms condition was close to significance,  $p = .056$ ,

while the difference between the AEI = 100 ms and the AEI = 300 ms condition was significant,  $p = .002$ . In Blocks 2, 3, and 4, there was no significant difference between the AEI = 0 ms and AEI = 100 ms conditions, all  $ps > .05$ , while both the AEI = 0 ms and the AEI = 100 ms conditions were significantly different from the AEI = 300 ms condition, all  $ps < .005$ . The main effect of block was not significant,  $F < 1$ .

In sum, the most important result regarding the analysis of temporal oculomotor control is that throughout all blocks the long effect delay (AEI = 300 ms) resulted in significantly longer saccade latencies than the short delay (AEI = 100 ms), and at least in three of the four blocks the long effect delay also resulted in significantly longer saccade latencies than the non-delayed (AEI = 0 ms) condition.

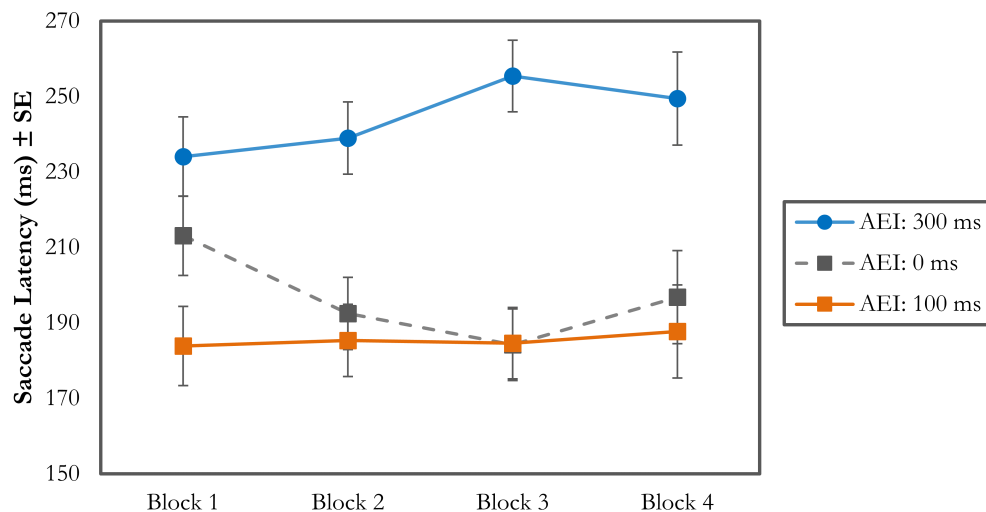


Figure 7. Mean saccade latencies (in ms) as a function of action-effect interval (AEI: 0 ms vs. 100 ms vs. 300 ms) and experimental block (1–4). Error bars represent standard errors of the mean ( $SE$  in ms). Figure adapted from Ref. [1].

### 3.1.3 DISCUSSION

The aim of the present study was to gain insight into the mechanisms underlying action-effect learning in the oculomotor domain. Accordingly, I analyzed spatial and temporal parameters of left/right eye movements that contingently triggered location-specific



---

changes in non-social, *inanimate* objects (traffic lights). Spatial aspects were addressed by examining spatial shifts of saccade target locations toward the anticipated location of the visual change, while temporal aspects were addressed by manipulating the temporal effect delay (AEI) and by studying the dynamics of action-effect learning over the course of the experiment.

The present results of Experiment 1 clearly support the hypothesis that action-effect anticipation in oculomotor control is reflected in corresponding spatial saccade characteristics in inanimate environments, similar to corresponding effects demonstrated in a social context (Herwig & Horstmann, 2011). Participants anticipated the specific change in the target objects and directed their initial saccade landing position toward the location of the anticipated effect (upper/lower lamp of traffic light). A shift in the saccade landing position was already present in the first block of trials (indicating fast buildup of action-effect associations) and across all AEI conditions. Thus, the upcoming effect clearly influences saccade planning. This is consistent with the assumption that spatial saccade preparation involves a mental representation of the upcoming effect, even in inanimate environments. Thus, these results – together with previous findings in a social context by Herwig and Horstmann (2011) – indicate that formation of oculomotor action-effect associations can occur in both types of environment (inanimate-nonsocial and animate-social).

Furthermore, the present findings extend previous, related observations. For example, Huestegge and Kreutzfeldt (2012) also reported evidence for action-effect anticipation in oculomotor control in an inanimate environment, but only with respect to congruency effects evident in saccade RTs, not in terms of spatial oculomotor control (i.e., saccade landing point). Conversely, another recent study already demonstrated effect anticipation evidenced by spatial oculomotor control (Pfeuffer et al., 2016). Crucially, in that study the spatial effects were contingent upon manual (not oculomotor) actions, which triggered spatially defined visual effects. Anticipation processes were evidenced by the occurrence of uninstructed

anticipatory saccades in direction of the subsequently occurring manual action effects. Despite this important difference to the present study, both studies confirm that spatial saccade parameters are a very sensitive (and probably more direct) measure of effect anticipation with many advantages over traditional, more indirect measures, like, for example, congruency effects reflected in RTs or error rate measures. These advantages include, *inter alia*, that anticipatory saccades can serve as a measure of anticipatory processes even in the absence of congruency effects in RTs or error rates. Further, measuring anticipatory saccades might be applied for the assessment of effect anticipation in populations where common RT-based paradigms reach their limits (e.g., clinical patients or young children, cf. discussion in Pfeuffer et al., 2016).

A second important result of the present study is the dependency of action latency on the temporal effect delay (action-effect interval) in the oculomotor domain. Especially in the AEI = 300 ms condition participants took more time to initiate their saccade toward one of the lateralized target objects compared to the AEI = 0 ms and 100 ms conditions. A similar observation was already reported for manual action control (Dignath et al., 2014), and was interpreted as evidence for the assumption that during the acquisition of action-effect associations actions become not only associated with the subsequent effect but also with the temporal AEI within an event file representation (Hommel et al., 2001). This temporal information is then assumed to be automatically retrieved during action selection, thus affecting latencies (see also Dignath & Janczyk, 2017). Note that in both previous studies on temporal AEI effects (Dignath et al., 2014; Dignath & Janczyk, 2017) the distance in length between short and long AEIs (50 ms vs. 2000 ms) was considerably more salient for the participant than in the present experiment, where the difference was rather subtle. Based on the fact that saccade latencies are usually shorter than corresponding manual RTs, it could be that temporal AEI effects scale accordingly. Of course, this proposal has to be empirically

---

addressed by using a broader spectrum of temporal intervals, before further conclusions can be made.

A further observation regarding the temporal AEI effect is that I could not find a clearly significant difference between the 0 ms and 100 ms AEI conditions. I suspect that the 100 ms spacing between these two conditions is too small to eventually affect saccade latencies. Again, this calls for future studies with a broader range of AEI intervals. Despite these limitations of the current data, the presented results are still in line with the claim of a general mechanism underlying action-effect learning by showing that effects of different temporal effect delays can also be observed in other than manual action domains, and that timing information regarding action effects plays a seminal role in the guidance of oculomotor actions.

Finally, a third important set of results is related to the dynamics of action-effect learning. Overall, the presented results indicate that saccade-effect associations can be easily acquired, since anticipation effects were already present from Block 1 (i.e., the rVLP averaged over the AEI conditions significantly differed from zero in all experimental blocks). This observation is in accordance with previous findings suggesting that only few learning trials are needed to acquire action-effect associations (Hommel, 2009; Pfeuffer et al., 2016). More interestingly, the results also showed that anticipation effects increased over the course of the experiment in the two delayed effect (AEI) conditions (see Herwig & Horstmann, 2011, for a corresponding finding in a setup using a constant delay of 100 ms). A plausible explanation for the result that action-effect learning only increased for the delayed effect conditions (but not in the immediate effect condition) might be that the visual change in the target object was less salient in the immediate condition, because the latter involves a display change during the saccade, which is less well perceived due to the well-known phenomenon of saccadic suppression. Thus, in the non-delayed condition, perception of the change must rely on comparing the fixated object with a memory representation of the object prior to its

fixation. This mechanism appears to be sufficient to explain a significant anticipation effect in the non-delayed condition already in Block 1. This memory-based source of information is principally also available in the delayed conditions, but here the change additionally provides a salient onset signal during fixation, which should attract even more attention (Enns, Austen, Di Lollo, Rauschenberger, & Yantis, 2001; von Mühlennen, Rempel, & Enns, 2005; Yantis & Jonides, 1984). Probably, this additional source of change information in the delayed conditions takes some time (over several blocks) to build up fully. On a general level, Experiment 1 further highlights the universality of the ideomotor view of action control (Greenwald, 1970; Herbart, 1825; James, 1890; Lotze, 1852) including its underlying principles and mechanisms by demonstrating its applicability not only in the manual (Elsner & Hommel, 2001; Kunde, 2001), but also in the oculomotor control domain (Herwig & Horstmann, 2011; Huestegge & Kreutzfeldt, 2012).

## 3.2 EXPERIMENTS 2–6

While Herwig and Horstmann (2011) focused on changes in *emotional* facial expression, the following experiments addressed *gaze responses* in virtual faces as action effects. This approach views sociomotor gaze control from another perspective that is relevant to a number of interaction situations in everyday life where the eye movements of an agent trigger gaze responses in the interaction partner. As the gaze-contingent stimulus change (i.e., as the effect of the participant's oculomotor action), I used faces with initially closed eyes that either focused their gaze toward or away from the participant (direct vs. averted gaze). The study was designed to address two main research questions. First, I tested whether learned associations between oculomotor actions and their effects on virtual interaction partners' gaze behavior determine saccade control (Experiment 2 and 3) by building on a well-established empirical setup for addressing anticipation-based oculomotor control from the (non-social)

---

domain of basic cognitive research (Huestegge & Kreutzfeldt, 2012). Specifically, participants responded to auditory stimuli (high-/low-pitched tones) with spatially corresponding saccades (upward/downward) toward one of two peripheral saccade targets. The targets were faces with closed eyes, which consistently changed as a result of the participant's gaze to either direct or averted gaze (i.e., the action effect), depending on the particular (upper vs. lower) face targeted by the participant. Using such an eye movement-contingent setup, participants should learn the perfect contingency between the face they look at (upper vs. lower face by initiating an upward vs. downward saccade, respectively) and the gaze response of that face (direct vs. averted gaze) over the course of the experiment, thus acquiring oculomotor action-effect associations. Since participants experienced the change of gaze direction only after carrying out a particular type of saccade on their own (i.e., either to the upper or to the lower face), one could conclude that in such a setting participants control the gaze of others (i.e., the gaze response of either of the two depicted faces) with their own gaze (i.e., to the upper or lower face).<sup>7</sup> To test whether saccade preparation toward the peripheral face was affected by the anticipation of the action's effect (i.e., the perceived direct or averted gaze), an additional task-irrelevant visual stimulus was centrally presented prior to the presentation of the imperative auditory stimulus (cf. Huestegge & Kreutzfeldt, 2012). This task-irrelevant stimulus (face with direct, averted, or no gaze) was either congruent, incongruent, or unrelated to the subsequent action effect. Following the ideomotor principle and findings from previous studies, saccade performance should depend on the congruency of the task-irrelevant visual stimulus with the saccade target, assuming the formation of effect representations prior to action initiation. That is, a congruent visual stimulus might serve as prime,

---

<sup>7</sup> A stronger, more natural form of controlling another's gaze with our own gaze would probably be a setting in which a participant freely chooses to move her/his eyes (instead of being instructed by means of an auditory imperative stimulus) in order to guide the gaze of another person into a certain direction or to a certain object. However, the present Experiments 2–6 are based on a previous established paradigm and designed to address anticipation-based oculomotor control under maximally controlled conditions.

which pre-activates the corresponding oculomotor response, eventually accelerating saccade initiation, while an incongruent visual stimulus might pre-activate the incorrect response, eventually slowing down saccade latencies. Thus, I expected to observe a congruency effect (using social gaze-related stimuli) as an index of effect anticipation. Corresponding evidence for a prediction of the other's gaze response to our own gaze would directly corroborate the idea that general ideomotor principles in oculomotor control can be extended to a social context, as posited by the sociomotor framework (Kunde et al., 2018).

By additionally manipulating the stimulus onset asynchrony (SOA) between the visual and auditory stimulus, I intended to explore whether the size of congruency effects may be affected by the time available for the build-up of effect anticipation (for similar manipulations see Huestegge & Kreutzfeldt, 2012, in the oculomotor, and Ziessler & Nattkemper, 2011, in the manual domain). Furthermore, I explored whether the anticipated effect signal (direct vs. averted gaze) influenced saccade latencies. Since in previous research saccade latencies were shown to be lower toward targets with high (vs. low) expected value (e.g., Milstein & Dorris, 2007), it is possible that saccade initiation is faster toward faces with direct gaze, since direct gaze represents a (potentially rewarding) valence-laden approach signal.

Second, I explored whether sociomotor gaze control differed from ideomotor gaze control without a social context. To do so, I also incorporated non-social control conditions in the present study (Experiments 4–6). Taking into account the peculiarities of social actions as outlined above, I explored whether social processing is qualitatively different from processing of non-social stimuli by testing whether anticipating action effects comprising social stimuli take longer to emerge (since social effects are often less immediate). Thus, congruency effects could be more pronounced for long versus short SOAs, especially in a social context.

---

### 3.2.1 METHODS

Five experiments were conducted, which involved different stimulus material but were otherwise identical. In Experiment 2, I used color photographs of faces in order to investigate gaze-interaction behavior in a realistic context. In Experiment 3, schematic black-and-white faces served as stimuli. Schematic faces are considered effective for the investigation of gaze-related behavior (Hietanen & Leppänen, 2003), most likely due to facilitated cue extraction based on visual simplicity. Experiments 4 to 6 involved stimuli that served as non-social control conditions. In Experiment 4, I used the inverted versions of the color photograph faces from Experiment 2 as stimuli. Inverted face photographs were classified as non-social stimuli based on previous research suggesting that processing of inverted faces lacks typical features associated with the holistic processing of upright faces (Vecera & Johnson, 1995; Yin, 1969). Inversion comes with the advantage of highly comparable visual input while altering the context from social to non-social processing. A second control condition (Experiment 5) was designed to conceptually replicate the findings of the original study (Huestegge & Kreutzfeldt, 2012), where geometric shapes served as stimuli. Viewed in conjunction with the elliptic face stimuli of Experiments 2 and 3, I opted for elliptic stimuli with a salient black square to ensure a certain degree of comparability between conditions. However, a final non-social control condition (Experiment 6) was designed to maximize the control of low-level differences between social and non-social stimulus conditions across experiments. Specifically, the stimuli of Experiment 6 differed from the schematic faces of Experiment 3 only with respect to the configuration of object features, while keeping overall visual low-level features constant.

#### 3.2.1.1 Participants

See *Table 1* for the full participant characteristics of each individual experiment. A total of 105 people took part in the experiments, but the data of four participants with an

unusually high error rate ( $> 30\%$  in at least one cell of the experimental design) had to be excluded. Additionally, the data of one participant was removed from analysis due to an error of the eye tracker. Thus, 20 participants took part in each of the five experiments and received either course credits or payment, resulting in 100 participants in total. A power analysis using G\*Power (Faul, Erdfelder, Lang, & Buchner, 2007) based on the large observed effect size of  $\eta^2_p = .43$  for the congruency effect in the precursor study (Huestegge & Kreutzfeldt, 2012) indicated a sample size of six participants as sufficient to observe a congruency effect with considerable probability (power = .95,  $\alpha = .05$ ). However, since I planned to compare data across experiments, I opted for an increased sample size ( $n = 20$ ) for each experiment to avoid an underpowered study. A sample size of 20 participants should ensure to observe a power of  $1 - \beta > .99$  to detect the original effect size of the congruency effect of  $\eta^2_p = .43$ , as reported in Huestegge and Kreutzfeldt (2012). All participants reported normal or corrected-to-normal vision and gave informed consent before completing the study. Participants were naïve with respect to the purpose of the experiment.

Table 1. Participant characteristics for each experiment

Experiment	Age range (in years)	Age (in years)		Gender distribution	
		<i>M</i>	<i>SD</i>	Women	Men
2: Upright faces	19 – 42	25.8	5.6	19	1
3: Schematic faces	19 – 35	24.5	4.3	17	3
4: Inverted faces	19 – 40	24.6	5.1	16	4
5: Abstract stimuli	20 – 34	24.0	3.6	16	4
6: Scrambled faces	19 – 29	23.6	2.8	16	4

Note. Values represent means (*M*) and standard deviations of the mean (*SD*). Table adapted from Ref. [2].

### 3.2.1.2 Apparatus

Note that the same technical equipment was utilized for Experiments 2–12. In a dimly lit room, participants were seated in front of a 20-in. cathode ray monitor (refresh rate: 100 Hz, spatial resolution:  $1024 \times 768$  pixels) with a viewing distance of 65 cm. Eye move-



ments of the right eye were recorded with a video-based eye tracker (EyeLink1000, SR Research, Ontario, Canada) at a sampling rate of 1000 Hz. A chin and a forehead rest stabilized the participant's head. The experiment was programmed using Experiment Builder (SR Research, Ontario, Canada). I relied on the EyeLink's build-in saccade parser (velocity threshold =  $30^\circ$ ; acceleration threshold =  $8000^\circ\text{s}^{-2}$ ) to categorize eye position data into meaningful events and states, like saccades, fixations, and blinks.

### 3.2.1.3 Stimuli

I have investigated action effects using different stimulus types in each experiment (see *Figure 8*). The social face stimuli comprised upright faces (color photographs of two female and two male faces; Experiment 2) and schematic face stimuli (black-and-white drawings reduced to the salient depiction of the main facial features; Experiment 3). The control stimuli were inverted faces (inverted versions of the color face photographs of Experiment 2; Experiment 4), abstract geometric stimuli (elliptic stimuli roughly comparable to the face shape, but without any strong resemblance to faces; Experiment 5), and scrambled face stimuli (black-and-white elliptic stimuli depicting the same elements as the schematic faces but arranged differently; Experiment 6).

The face stimuli of Experiments 2–4 were available in three variants, reflecting different gaze conditions: direct gaze (i.e., gazing toward the observer), averted gaze (i.e., gazing to the left/right away from the observer) and closed eyes (baseline). Note that there is no real “gaze” visible in Experiments 5 (abstract geometrical stimuli) and 6 (scrambled face stimuli), which were elliptical white shapes with one (Experiment 5) or two (Experiment 6) black squares to simulate the different gaze variations (see *Figure 8*). The abstract geometric stimuli (Experiment 5) were designed in a way that they are roughly comparable to the face stimuli (with respect to overall shape and size), but involve a slightly more salient effect in terms of reduced visual features, a larger black square and a more pronounced displacement of the square. The different gaze variations were realized by an ellipsis with a central black

square as an abstract, non-social equivalent to direct gaze, by a lateral black square (to the left/right) as an equivalent to averted gaze, and with the absence of any black square indicating baseline condition. The scrambled face stimuli (Experiment 6), however, were designed in a way that they are maximally comparable to the schematic face stimuli with respect to overall shape, size, and amount of visual information, i.e., using the same pixel information: Each facial feature of the schematic face was used in the scrambled face stimulus. As an exception, the black pupil circles of the schematic faces were adapted to black squares to avoid obvious eye gaze associations. To ensure high comparability between the schematic face and scrambled face stimuli, the principle of symmetry as implemented in the schematic face stimuli was transferred to the control stimuli. Analogous to the different gaze conditions in the face stimuli, the elliptic shape with two central black squares (arranged vertically) was considered a non-social equivalent to direct gaze, the black squares placed laterally (to the left/right) resembled averted gaze, and the absence of any black square served as a baseline condition. Importantly, however, the abstract and control stimuli should not elicit any obvious face/gaze association.

The face stimuli (Experiment 2–4) were of emotionally neutral facial expression. The photographic faces (Experiment 2 and 4) were presented in vertical ellipses with a size of  $4.6^\circ \times 6.0^\circ$  of visual angle (maximum width  $\times$  maximum height), horizontally extending from ear to ear and vertically from chin to hairline. The size of the schematic face and abstract geometric stimulus ellipses corresponded to the size of the actual oval face shape of the photographic faces (ears and hairs excluded), resulting in a slightly smaller ellipse size ( $3.5^\circ \times 5.3^\circ$  of visual angle).

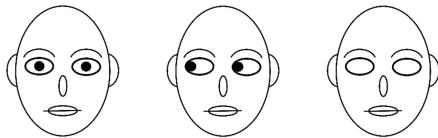
The upright and inverted face stimuli were photographs taken by myself and processed using GIMP (version 2.8). For each of the four face models, we used the closed eye photograph as a template and inserted the eye area from the remaining gaze conditions (i.e., direct/left averted/right averted gaze). For the inverted face condition, the face stimuli were

rotated by 180°. A high- and a low-pitched tone (50 ms beep of 1000 or 400 Hz) served as acoustic stimuli.

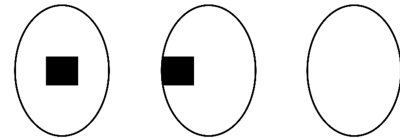
#### A: Experiment 2



#### B: Experiment 3



#### D: Experiment 5



#### C: Experiment 4

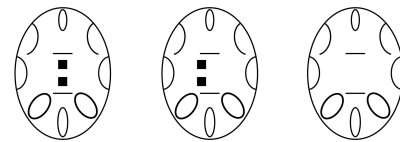


Direct  
gaze

Averted  
gaze

Closed  
eyes  
(Baseline)

#### E: Experiment 6



Central  
effect

Lateral  
effect

No effect  
(Baseline)

*Figure 8.* Stimulus types used in Experiment 2 (A; upright face stimuli), Experiment 3 (B; schematic face stimuli), Experiment 4 (C; inverted face stimuli), Experiment 5 (D; abstract geometric stimuli), and Experiment 6 (E; scrambled face stimuli). For each experiment, three different variants are shown (from left to right: direct gaze vs. averted gaze vs. closed eyes (baseline) for Experiments 2–4, and central square(s) vs. lateral square(s) vs. no square for the corresponding variants for Experiments 5 and 6). Figure adapted from Ref. [2].

#### 3.2.1.4 Procedure

The experimental procedure (which was highly similar to the procedure used in Huestegge & Kreutzfeldt, 2012) was the same for each experiment, the only crucial difference was the stimulus material. *Figure 9* depicts the schematic trial sequence. First, participants were familiarized with the acoustic stimuli used during the experiment. A black screen with a white central fixation cross (0.4° of visual angle) was presented. Two baseline stimuli were presented in the periphery at an eccentricity of 7.9° of visual angle above and below the fixation cross, serving as saccade targets. Then, the central fixation cross was substituted by

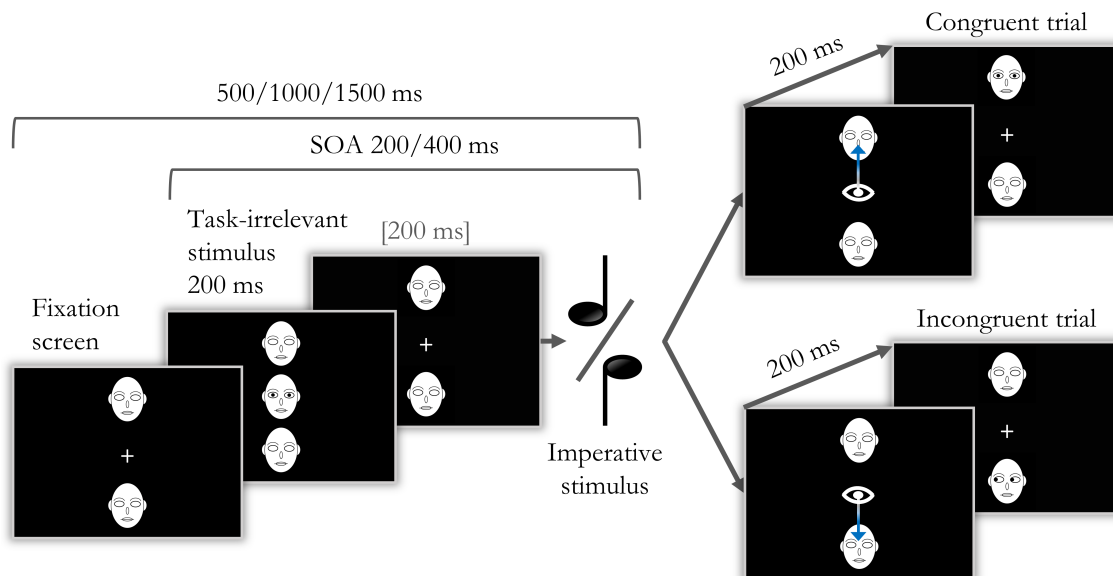
a task-irrelevant visual stimulus, presented for 200 ms. The visual stimulus was either a face stimulus with direct gaze, averted gaze, or closed eyes (Experiments 2–4), or a stimulus with central square(s), lateral square(s), or no square(s) (Experiments 5 and 6), with all variants equally distributed across trials. Crucially, the visual stimulus was either congruent or incongruent (for direct/averted gaze stimuli and central/lateral square stimuli), or unrelated (for baseline stimuli) to the subsequently occurring effect stimulus. Then, the imperative auditory stimulus of either high or low pitch was presented. The time interval between trial onset and the onset of the imperative auditory stimulus was randomized (500/1000/1500 ms, equally distributed). The duration of the interval between the visual and the auditory stimulus (SOA) amounted to 200 or 400 ms. In the case of the latter, the fixation cross reappeared for the additional 200 ms time period between offset of the visual stimulus and onset of the auditory stimulus and was visible until the end of the trial. In case of the SOA = 200 ms condition, the central fixation cross reappeared with onset of the auditory stimulus and was also presented until the end of the trial.

Participants were instructed to shift their gaze from central fixation as quickly as possible upward or downward in accordance with the pitch of the auditory stimulus toward one of the peripheral saccade targets and to come back to the fixation cross immediately afterwards. Contingent upon the participant's gaze shift toward the upper or lower target area, the baseline stimulus in the target area changed 200 ms after the saccade entered the target area: For example, one half of the participants consistently experienced the onset of an averted gaze when looking at the upper face and the onset of a direct gaze when looking at the lower face (or the corresponding equivalents for abstract stimuli). This assignment was reversed for the other half of participants. The direction of the averted gaze or lateral square stimulus (i.e., averted gaze/lateral square to the left/right) was counterbalanced. After the return saccade to the central fixation cross, the baseline stimuli reappeared at the target locations in the periphery. In case of an erroneous saccade response, for example, when a high

---

tone had indicated to look toward the upper target area but the participant looked toward the lower target area, no error feedback appeared. Moreover, a saccade toward the upper or lower position *always* triggered the associated effect stimulus (direct vs. averted gaze or central vs. lateral square) irrespective of whether participants acted in accordance with the pitch of the imperative tone or not. Thus, gaze direction of the effect stimulus (or the corresponding equivalent for abstract stimuli) was entirely determined by the (incorrect) action participants actually carried out rather than by the action they should have carried out. In this way, the triggered effect was contingent on participants' action rather than on the preceding tone. However, these erroneous trials were excluded from the analysis of saccade latencies. Please note that the unrelated task-irrelevant visual stimulus (face with closed eyes in Experiments 2–4 or ellipsis without any black squares in Experiments 5 and 6) represented a neutral baseline with respect to the congruency between the task-irrelevant visual stimulus and the effect stimulus in that it never served as an effect stimulus. The next trial started 1500 ms after the imperative auditory stimulus. If no saccade to one of the two peripheral target areas was detected within 1500 ms, a feedback message appeared displaying “*Keine Reaktion*” (German for “no response”), and the next trial started. Note that in Experiments 2 and 4 (upright and inverted face stimuli), only stimuli of the same model face were presented within a trial.

Each of the experiments (around 40 min.) comprised 16 experimental blocks with 36 trials each (576 trials in total). Prior to each block, a calibration of the eye tracker was performed. Trial order was pseudo-randomized (maximum sequence length of the same auditory stimulus was restricted to eight).



*Figure 9.* Trial sequence depicted for schematic face stimuli: After the presentation of a black screen with a white fixation cross and two peripheral baseline face stimuli (without gaze), the fixation cross was replaced by a task-irrelevant stimulus, which was either a baseline face (without any visible gaze) stimulus or a face with direct or averted gaze. Then, an imperative auditory stimulus (high vs. low) signaled the participant to respond with a corresponding gaze shift (up vs. down). As an effect of the saccade toward the upper or lower face, either a direct or averted gaze (contingent upon the targeted face) was displayed. Note that the task-irrelevant stimulus was either congruent, incongruent, or unrelated to the subsequent gaze response of the targeted face. The trial sequence shown in this figure depicts a congruent (top right) and incongruent (bottom right) trial sequence for a direct gaze as the task-irrelevant stimulus. In case of an averted gaze as task-irrelevant stimulus (not depicted in the present figure), a gaze shift toward the upper target face (triggering direct gaze) would represent an incongruent trial, and a gaze shift toward the lower target face (triggering averted gaze) would represent a congruent trial. For the unrelated congruency condition, the task-irrelevant stimulus was a baseline face without any visible gaze. Figure adapted from Ref. [2].

### 3.2.1.5 Design and analysis

I analyzed saccade latency as well as the direction of saccade responses (up vs. down). Saccade latency was defined as the interval between the onset of the imperative auditory stimulus and the initiation of a saccade with an amplitude of at least one third of the distance between the fixation cross and the center of the saccade target. Oculomotor responses into the wrong direction were counted as errors. Mean saccade latencies (i.e., oculomotor RTs), were analyzed for correct trials only.<sup>8</sup> Each experiment involved three within-subjects factors: *congruency* (congruent vs. incongruent vs. unrelated), *SOA* (200 ms vs. 400

<sup>8</sup> Note that this procedure applies to all subsequent experiments within this work, unless stated otherwise.

---

ms), and *direction* of the action effect (direct gaze vs. averted gaze in Experiments 2–4, and central vs. lateral effect in Experiment 5 and 6).

Between-experiment comparisons served to assess any potential difference between a social versus non-social context. We conducted subsequent repeated-measures ANOVAs with the between-subjects factor experiment and the within-subjects factors congruency (congruent vs. incongruent vs. unrelated) and SOA (200 ms vs. 400 ms), and direction of the action effect (direct gaze vs. averted gaze in Experiments 2–4, and central vs. lateral effect in Experiment 5 and 6). We compared the results of Experiment 2 and Experiment 4 (upright and inverted face stimuli), and Experiment 3 and 6 (schematic face and scrambled face stimuli), since corresponding stimulus material was perfectly comparable across contexts in this case. We refrained from a statistical comparison of experiments with photographic and non-photographic stimuli since behavioral differences cannot be attributed unambiguously to the context variation per se but might instead relate to low-level differences in stimulus material. Data and analysis scripts are available on the Open Science Framework (<https://osf.io/q5p3k/>).

### 3.2.2 RESULTS

Prior to the analysis I removed all trials with tracking errors as well as trials involving blinks, anticipatory saccades (latency < 60 ms), saccades to the left or right (instead of up/down saccades toward the targets), or the absence of any saccade meeting the response criteria outlined above. This cleansing procedure yielded 386 trials of Experiment 2 to be excluded, which corresponded to 3.4% of the total number of trials (20 participants with 576 each resulting in 11520 trials in total). In Experiment 3, 572 trials (5.0%) had to be excluded, 476 trials (4.1%) in Experiment 4, 362 trials (3.1%) in Experiment 5, and 421 trials (3.7%) in Experiment 6. RTs of trials with correct responses and error rates were submitted to a three-

way ANOVA with congruency, SOA and direction as within-subjects factors for each experiment. Note that we report the results of each experiment separately for temporal oculomotor performance (saccade RTs) and oculomotor performance in terms of error rates by concentrating on commonalities across experiments first before reporting differences.

### 3.2.2.1 Temporal oculomotor performance

*Table 2* depicts an overview of the mean saccade RTs (overall mean and the mean for each congruency condition) for each experiment. In all experiments, the main effect of SOA reached significance, indicating longer RTs for the SOA = 200 ms condition compared to the SOA = 400 ms condition. Specifically, RTs for the short versus long SOA condition amounted to 281 ms ( $SE = 12$  ms) versus 252 ms ( $SE = 11$  ms) in Experiment 2,  $F(1, 19) = 31.35, p < .001, \eta^2_p = .62$ , and to 264 ms ( $SE = 9$  ms) versus 235 ms ( $SE = 9$  ms) in Experiment 3,  $F(1, 19) = 80.12, p < .001, \eta^2_p = .81$ . In Experiment 4, RTs for the short versus long SOA condition were 267 ms ( $SE = 11$  ms) versus 245 ms ( $SE = 11$  ms),  $F(1, 19) = 54.69, p < .001, \eta^2_p = .74$ , and 294 ms ( $SE = 15$  ms) versus 265 ms ( $SE = 13$  ms) in Experiment 5,  $F(1, 19) = 92.53, p < .001, \eta^2_p = .83$ . In Experiment 6, RTs for the short SOA amounted to 296 ms ( $SE = 13$  ms) versus 263 ms ( $SE = 11$  ms) in the long SOA condition,  $F(1, 19) = 92.39, p < .001, \eta^2_p = .83$  (see *Figure 10*).

In Experiment 2 (upright face stimuli), there was no significant effect of congruency,  $F < 1$ . Additionally, none of the remaining effects reached statistical significance, neither the effect of direction nor the interaction of congruency and direction or the interaction of congruency and SOA (all  $F$ s  $< 1$ ). Further, neither the interaction of direction and SOA,  $F(1, 19) = 1.28, p = .272, \eta^2_p = .06$ , nor the three-way interaction,  $F < 1$ , were significant.

By contrast, the effect of congruency reached significance in Experiment 3 (schematic face stimuli),  $F(2, 38) = 3.75, p = .033, \eta^2_p = .17$ . RTs were shortest in congruent trials, longest in incongruent trials, and at an intermediate level in trials with unrelated visual stimuli



(see *Table 2*). Pairwise comparisons showed a significant difference in RTs between the congruent and the incongruent condition,  $p = .016$ , while there was no significant difference between congruent and unrelated trials,  $p = .287$ , and between incongruent and unrelated trials,  $p = .063$ . The interaction of direction and congruency,  $F(2, 38) = 3.44$ ,  $\varepsilon = .74$ ,  $p = .059$ ,  $\eta^2_p = .15$ , was not significant. The main effect of direction was not significant,  $F < 1$ . None of the remaining interactions reached significance, neither the interaction of congruency and SOA,  $F(2, 38) = 2.41$ ,  $p = .104$ ,  $\eta^2_p = .11$ , nor the interaction of direction and SOA,  $F(1, 19) = 2.03$ ,  $p = .171$ ,  $\eta^2_p = .10$ , or the three-way interaction,  $F < 1$ .

In Experiment 4 (inverted face stimuli), no significant main effect of congruency occurred,  $F(2, 38) = 2.69$ ,  $\varepsilon = .70$ ,  $p = .101$ ,  $\eta^2_p = .12$ . Furthermore, none of the remaining effects was significant, neither the main effect of direction,  $F < 1$ , nor the interaction of direction and SOA,  $F(1, 19) = 1.78$ ,  $p = .198$ ,  $\eta^2_p = .09$ , or the interactions of direction and congruency, SOA and congruency, and the three-way interaction (all  $F$ s  $< 1$ ).

In Experiment 5 (abstract geometric stimuli), also a significant main effect of congruency could be observed,  $F(2, 38) = 7.11$ ,  $p = .002$ ,  $\eta^2_p = .27$ , indicating shorter RTs in the congruent condition compared to the incongruent condition,  $p = .003$  for the two-tailed pairwise  $t$ -test. RTs were also significantly smaller for trials with unrelated stimuli compared to incongruent trials,  $p = .021$ , but there was no difference between trials with congruent and unrelated stimuli,  $p = .145$ . There was a significant interaction of direction and SOA,  $F(1, 19) = 8.08$ ,  $p = .010$ ,  $\eta^2_p = .30$ . Note that we refrain from interpreting this interaction as the direction manipulation in Experiment 5 was only implemented to ensure comparability with the face stimuli, but without any reasoning about why central vs. lateral effects should elicit different saccade latencies. Further, there was no significant main effect of direction,  $F(1, 19) = 1.17$ ,  $p = .293$ ,  $\eta^2_p = .06$ , no significant interaction of direction and congruency,  $F(2, 38)$

$= 1.83, p = .175, \eta^2_p = .09$ , no significant interaction of congruency and SOA, and finally no significant three-way interaction (both  $F_s < 1$ ).

There was an influence of congruency on saccade RTs in Experiment 6 (scrambled face stimuli),  $F(2, 38) = 9.11, p = .001, \eta^2_p = .32$ . However, this effect was not due to a significant difference between congruent and incongruent trials,  $p = .476$ , but due to significantly higher RTs for unrelated trials compared to both, congruent,  $p = .002$ , and incongruent trials,  $p = .003$ . None of the remaining effects reached significance (all  $F_s < 1$ ).

To explore whether the congruency effect remained stable over the course of the experiment, we extended the analysis and conducted a repeated-measures ANOVA with the additional within-subjects factor *experiment half* (first half vs. second half). I will first report the stability analysis for Experiments 3 and 5, where congruency effects were actually observed, and only report effects of experiment half as well as effects involving both the factors congruency and experiment half here. For schematic face stimuli (Experiment 3), we observed a significant effect of experiment half, indicating that saccade latencies decreased from the first to the second half of the experiment,  $F(1, 19) = 10.32, p = .005, \eta^2_p = .35$ . None of the remaining relevant effects were significant (all  $F_s < 1$ , except for the three-way interaction of congruency, SOA, and experiment half,  $F(2, 38) = 2.37, \varepsilon = .77, p = .123, \eta^2_p = .11$ ). In Experiment 5 (abstract geometric stimuli) the effect of experiment half was not significant,  $F(1, 19) = 3.51, p = .077, \eta^2_p = .16$ . Also, none of the other relevant interactions revealed significant effects, neither the interaction of congruency and experiment half, nor the interaction of congruency, SOA, and block (both  $F_s < 1$ ), the interaction of congruency, direction, and experiment half,  $F(2, 38) = 2.77, p = .075, \eta^2_p = .13$ , or the four-way interaction,  $F(2, 38) = 2.32, p = .112, \eta^2_p = .11$ .

For the sake of completeness, I also report the respective analyses for Experiments 2, 4 and 6. In Experiment 2 (upright face stimuli), neither the effect of experiment half,  $F(1,$

19) = 2.57,  $p = .125$ ,  $\eta^2_p = .12$ , nor the three-way interaction of congruency, SOA, and experiment half,  $F(2, 38) = 1.28$ ,  $p = .290$ ,  $\eta^2_p = .06$ , or any other relevant interaction revealed significant results (all  $F$ s < 1). In Experiment 4 (inverted face stimuli) participants responded faster in the first versus second half of the experiment,  $F(1, 19) = 11.77$ ,  $p = .003$ ,  $\eta^2_p = .38$ . None of the other relevant interactions were significant, neither the interaction of congruency and experiment half,  $F(2, 38) = 1.85$ ,  $p = .172$ ,  $\eta^2_p = .08$ , nor the three-way interaction of congruency, direction, and experiment half,  $F(2, 38) = 1.66$ ,  $p = .204$ ,  $\eta^2_p = .08$ , or the interaction of congruency, SOA, and experiment half,  $F(2, 38) = 2.36$ ,  $p = .108$ ,  $\eta^2_p = .11$ . The four-way interaction was not significant,  $F < 1$ . In Experiment 6 (scrambled face stimuli) no effect of experiment half was present,  $F(1, 19) = 2.30$ ,  $p = .146$ ,  $\eta^2_p = .11$ . The four-way interaction was significant,  $F(2, 38) = 3.54$ ,  $p = .039$ ,  $\eta^2_p = .16$ . None of the remaining relevant interaction effects was significant (all  $F$ s < 1, except for the three-way interaction of congruency, SOA, and experiment half,  $F(2, 38) = 2.96$ ,  $\varepsilon = .69$ ,  $p = .086$ ,  $\eta^2_p = .14$ ).

Finally, I explored whether the direction of the task-irrelevant stimulus (direct/averted gaze direction in Experiment 3; central/lateral effect direction in Experiment 5) that was presented shortly before the imperative auditory stimulus affected saccade latencies. To do so, I conducted a two-way repeated-measures ANOVA involving the within-subjects factors *task-irrelevant stimulus* (direct gaze vs. averted gaze direction in Experiment 3, and central vs. lateral effect direction in Experiment 5) and *congruency* (congruent vs. incongruent) as well as the between-subjects factor *experiment* (Experiment 3 vs. Experiment 5). The unrelated baseline condition was excluded from this analysis, and thus only effects involving the factor task-irrelevant stimulus are reported. Saccade latencies were reduced for around 5 ms for a task-irrelevant stimulus displaying direct gaze/central effect direction as compared to a task-irrelevant stimulus displaying averted gaze/lateral effect direction,  $F(1, 38) = 12.39$ ,  $p = .001$ ,  $\eta^2_p = .25$ . This effect of task-irrelevant stimulus was not affected by

any of the other factors (all  $F_s < 1$  for the statistical comparisons involving the factor task-irrelevant stimulus).<sup>9</sup>

Taken together, the results of the temporal oculomotor performance analysis showed congruency effects (in terms of significant differences between congruent and incongruent conditions) for both social and non-social stimuli (see *Figure 10*). However, the effect only occurred for non-photographic stimuli (i.e., for schematic face stimuli and for abstract stimuli) where the visual action effect was sufficiently salient either due to embedding the effect into a schematic face context (Experiment 3) or due to the isolated presentation of the effect (Experiment 5). In these two experiments, I observed an adverse effect of effect incongruency rather than a beneficial effect of effect congruency (when compared to the baseline conditions). This was different when the black-and-white scrambled faces served as stimuli (Experiment 6), where RTs were reduced for both congruent and incongruent trials compared to baseline. Consistently across all experiments, saccades were initiated faster with longer SOA. Moreover, the congruency effect as observed in Experiments 3 and 5 remained stable and was evident in saccadic RTs not only in the first half, but also in the second half of the experiment.

---

<sup>9</sup> Please note that none of the statistical comparisons revealed significant results (all  $F_s < 1$ ) when conducting a two-way repeated-measures ANOVA with the factors *task-irrelevant stimulus* (direct gaze vs. averted gaze direction in Experiments 2 and 4; central vs. lateral effect direction in Experiment 6) and *congruency* (congruent vs. incongruent) separately for Experiments 2, 4, and 6.

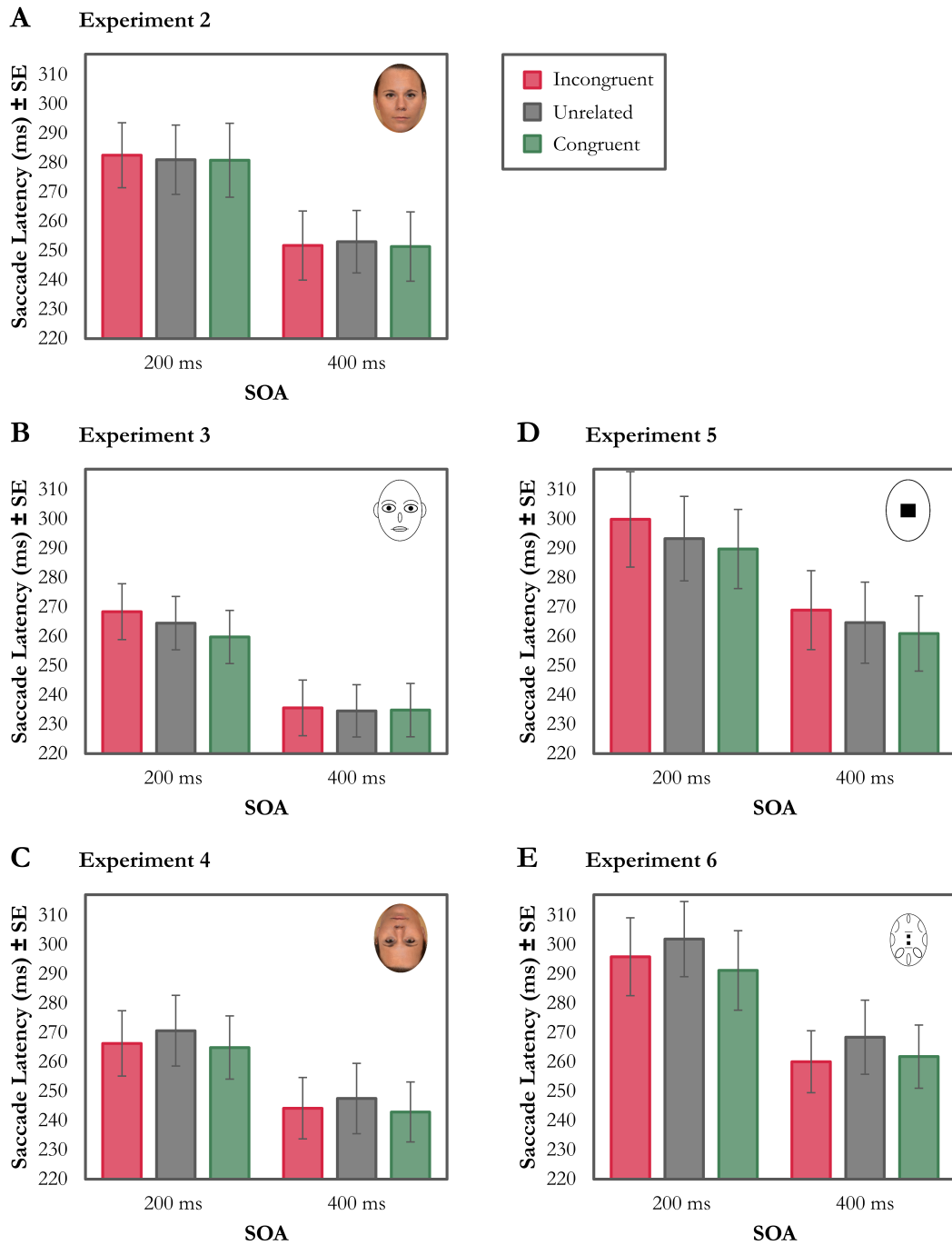


Figure 10. Saccade latencies (in ms) as a function of congruency (congruent vs. incongruent vs. unrelated) and stimulus onset asynchrony (SOA; 200 ms vs. 400 ms) in Experiment 2 (A; upright face stimuli), Experiment 3 (B; schematic face stimuli), Experiment 4 (C; inverted face stimuli), Experiment 5 (D; abstract geometric stimuli), and Experiment 6 (E; scrambled face stimuli). Error bars represent standard errors of the mean ( $SE$  in ms). Figure adapted from Ref. [2].

Table 2 Mean performance parameters (saccade latency and error rate) across experiments

Experiment	Mean saccade latency		Mean error rate	
	<i>M</i> (ms)	<i>SE</i>	<i>M</i> (%)	<i>SE</i>
<b>2: Upright face stimuli</b>	<b>267</b>	<b>11</b>	<b>8.2</b>	<b>1.3</b>
Congruent	266	12	7.9	1.3
Incongruent	267	11	8.6	1.5
Unrelated	267	11	8.1	1.3
<b>3: Schematic face stimuli</b>	<b>250</b>	<b>9</b>	<b>7.0</b>	<b>1.1</b>
Congruent	247	9	6.6	1.1
Incongruent	252	9	7.8	1.1
Unrelated	250	9	6.6	1.0
<b>4: Inverted face stimuli</b>	<b>256</b>	<b>11</b>	<b>7.8</b>	<b>1.0</b>
Congruent	254	10	7.8	1.1
Incongruent	255	11	7.7	1.2
Unrelated	259	11	7.9	1.0
<b>5: Abstract geometric stimuli</b>	<b>280</b>	<b>14</b>	<b>5.7</b>	<b>0.8</b>
Congruent	275	13	4.8	0.9
Incongruent	284	15	6.6	0.8
Unrelated	279	14	5.6	0.8
<b>6: Scrambled face stimuli</b>	<b>280</b>	<b>12</b>	<b>7.0</b>	<b>1.1</b>
Congruent	277	12	7.1	1.2
Incongruent	278	12	7.2	1.1
Unrelated	285	13	6.7	1.0

Note. Means (*M*) and standard errors (*SE*) are presented. Table adapted from Ref. [2].

### 3.2.2.2 Oculomotor error rates

Mean error rates for each experiment and each congruency condition are listed in Table 2. Across all experiments, we observed a significant main effect of SOA, indicating fewer error rates in the short SOA = 200 ms condition compared to the long SOA = 400 ms condition. Specifically, the error rate amounted to 6.8% (*SE* = 1.2%) versus 9.6% (*SE* = 1.7%) in the short versus long SOA condition in Experiment 2,  $F(1, 19) = 6.57, p = .019, \eta^2_p = .26$ , and to 5.5% (*SE* = 1.0%) versus 8.5% (*SE* = 1.3%) in the short versus long SOA condition in Experiment 3,  $F(1, 19) = 10.67, p = .004, \eta^2_p = .36$ . In Experiment 4, error rates

for the short versus long SOA condition were 6.5% ( $SE = 0.8\%$ ) versus 9.1% ( $SE = 1.4\%$ ),  $F(1, 19) = 8.67, p = .008, \eta^2_p = .31$ , and to 4.9% ( $SE = 0.6\%$ ) versus 6.4% ( $SE = 1.0\%$ ) in Experiment 5,  $F(1, 19) = 5.94, p = .025, \eta^2_p = .24$ . In Experiment 6, error rates amounted to 6.2% ( $SE = 1.0\%$ ) versus 7.8% ( $SE = 1.2\%$ ) in the short versus long SOA condition,  $F(1, 19) = 5.46, p = .031, \eta^2_p = .23$ .

In Experiment 2 (upright face stimuli), error rates were not significantly affected by the congruency or direction manipulation (both  $F_s < 1$ ). Further, the interaction of congruency and direction,  $F(2, 38) = 1.64, p = .207, \eta^2_p = .08$ , and of congruency and SOA,  $F < 1$ , did not yield any significant effects, as well as the interaction of direction and SOA,  $F(1, 19) = 3.52, p = .076, \eta^2_p = .16$ , and the three-way interaction,  $F(2, 38) = 2.58, p = .089, \eta^2_p = .12$ .

Similarly, congruency did not affect error rates in Experiment 3 (schematic face stimuli),  $F(2, 38) = 2.61, p = .087, \eta^2_p = .12$ . Neither the interaction of congruency and SOA,  $F(2, 38) = 2.84, p = .071, \eta^2_p = .13$ , nor the interaction of congruency and direction,  $F(2, 38) = 1.62, \varepsilon = .76, p = .217, \eta^2_p = .08$ , were significant. None of the remaining effects were significant (all  $F_s < 1$ ).

In Experiment 4 (inverted face stimuli), we observed no other significant effects except for the main effect of SOA as reported above (all  $F_s < 1$ , except for the three-way interaction,  $F(2, 38) = 1.05, p = .360, \eta^2_p = .05$ ).

In Experiment 5 (abstract geometric stimuli), however, we observed a main effect of congruency,  $F(2, 38) = 3.93, p = .028, \eta^2_p = .17$ . Participants committed fewer errors in congruent trials compared to incongruent trials,  $p = .024$ , while there was no significant difference between congruent and unrelated,  $p = .103$ , and unrelated and incongruent trials,  $p = .171$ . The three-way interaction,  $F(2, 38) = 3.90, p = .029, \eta^2_p = .17$ , was significant. Follow-up analyses showed no significant influence of SOA,  $F(1, 19) = 2.74, p = .114, \eta^2_p = .13$ , of

congruency,  $F(2, 38) = 2.34, p = .110, \eta^2_p = .11$ , or the interaction of both,  $F(2, 38) = 2.30, p = .114, \eta^2_p = .11$ , for central effect stimuli. For lateral effect stimuli, we observed a significant effect of SOA,  $F(1, 19) = 4.46, p = .048, \eta^2_p = .19$ , but no congruency effect,  $F(2, 38) = 2.84, p = .071, \eta^2_p = .13$ , and a non-significant interaction of both,  $F(2, 38) = 2.47, p = .098, \eta^2_p = .12$ . No other statistical comparison was significant, neither the effect of direction,  $F(1, 19) = 2.09, p = .165, \eta^2_p = .1$ , nor any other two-way interaction (all  $F$ s  $< 1$ ).

The main effect of congruency was not significant in Experiment 6 (scrambled face stimuli),  $F < 1$ . However, I observed a significant interaction of congruency and direction,  $F(2, 38) = 3.51, p = .040, \eta^2_p = .16$ . Separate one-way ANOVAs revealed a significant influence of congruency for the central effect condition,  $F(2, 38) = 3.42, p = .043, \eta^2_p = .15$ , but not for the lateral effect condition,  $F(2, 38) = 1.08, p = .350, \eta^2_p = .05$ . In the central effect direction, participants committed fewer errors in unrelated versus incongruent trials,  $p = .026$ , while there was no significant difference between congruent and incongruent,  $p = .060$ , or between congruent and unrelated trials,  $p = .989$ . None of the remaining effects approached significance, neither the effect of direction  $F < 1$ , the interaction of direction and SOA,  $F(1, 19) = 2.00, p = .174, \eta^2_p = .10$ , the interaction of congruency and SOA,  $F(2, 38) = 2.23, p = .121, \eta^2_p = .11$ , nor the three-way interaction,  $F < 1$ .

Evidence for congruency effects in error rates was observed for the abstract non-social stimuli. There were no (or only very weak indications of) congruency effects for the schematic faces, the photographic (upright and inverted) and the non-social control stimuli. In sum, the error rate analysis showed that error rates can serve as an indicator of anticipation-related congruency effects, albeit in a somewhat weaker form as compared to saccade latencies. More importantly, participants committed more errors in incongruent compared to congruent trials in Experiments 3 (as reflected in the descriptive values) and Experiment 5 (as reflected in the significant effect of congruency). This leads to the conclusion that



speed-accuracy tradeoffs do not compromise the interpretation of the congruency effects that emerged in RTs in these two experiments since the error rate pattern reflects the performance pattern found in saccade latencies (faster RTs in congruent vs. incongruent trials).

Interestingly, participants committed fewer errors in the short (vs. long) SOA condition across all experiments. When viewed in conjunction with the higher saccade latencies for the short SOA conditions, this pattern points toward a speed-accuracy tradeoff: When participants respond particularly fast (in the long SOA condition), these fast responses tend to be slightly more error-prone. When compared to experiments using a similar setting (e.g., Huestegge & Kreutzfeldt, 2012), error rates are generally high in the current reported experiments, especially in Experiments 2 and 4. This discrepancy in error rates between experiments using photographic versus non-photographic stimuli is probably due to the utilization of visually more complex photographic stimuli, potentially eliciting more visual distraction as compared to schematic faces (Experiment 3) and abstract geometric stimuli (Experiment 5).

*Between-experiment comparisons<sup>10</sup>*

The between-experiment comparison of Experiment 2 and Experiment 4 (upright vs. inverted face stimuli) did not reveal any significant effects involving the factor experiment (all  $F$ s < 1, except for the interaction of congruency and experiment,  $F(2, 76) = 1.19, p = .309, \eta^2_p = .03$ , the interaction of SOA and experiment,  $F(1, 38) = 1.33, p = .256, \eta^2_p = .03$ , and the four-way interaction,  $F(2, 76) = 1.26, p = .289, \eta^2_p = .03$ ). The comparison of Experiments 3 and 6 (schematic face and scrambled face stimuli) revealed a significant interaction of congruency and experiment,  $F(2, 76) = 6.37, p = .003, \eta^2_p = .14$ , indicating that congruency affected saccade latencies differently in Experiments 3 and 6. While I observed

---

<sup>10</sup> Note that only statistical comparisons involving the factor experiment were reported for the between-experiment comparisons.

a significant difference between the congruent and incongruent condition in Experiment 3, this was not the case for Experiment 6 (where congruent and incongruent conditions were comparable but both significantly different from the unrelated condition). The interaction of SOA and experiment was not significant,  $F < 1$ , suggesting that the long SOA reduced saccade latencies similarly across experiments. In terms of absolute numbers, saccade latencies were smaller for schematic face stimuli compared to control stimuli, but the statistical comparison failed to reach significance,  $F(1, 38) = 4.06, p = .051, \eta^2_p = .10$ . The three-way interaction of direction, congruency, and experiment was not significant,  $F(2, 76) = 1.53, p = .224, \eta^2_p = .04$ . None of the remaining interactions involving the factor experiment was significant (all  $F_s < 1$ ).

Due to the significant congruency effects in Experiments 3 and 5 (schematic face and abstract geometric stimuli), I ran a post-hoc between-experiment comparison including these two experiments to assess the influence of a social versus non-social context, even though stimulus material was not perfectly comparable across contexts, but similar in terms of a reduction to black-and-white (as opposed to photographic) stimuli. The interaction of congruency and experiment was not significant,  $F(2, 76) = 1.14, p = .326, \eta^2_p = .03$ . This finding indicates that congruency did not affect saccade latencies differently across these two experiments. In terms of absolute numbers, saccade latencies were slightly lower for schematic face versus abstract stimuli, but the statistical comparison did not reach significance,  $F(1, 38) = 3.30, p = .077, \eta^2_p = .08$ . None of the remaining effects involving the factor experiment returned significant results (all  $F_s < 1$ , except for the interaction of direction and experiment,  $F(1, 38) = 1.22, p = .276, \eta^2_p = .03$ ).

In sum, these subsequent between-experiment comparisons demonstrated that the pattern of the congruency effect as observed in Experiment 3 (schematic face stimuli) versus Experiment 6 (scrambled face stimuli) is significantly different. While congruency between

---

the visual stimulus and the subsequent action effect facilitated a saccadic response in Experiment 3, I did not observe such a facilitating effect in Experiment 6, where the congruent and incongruent conditions did not differ. Further, the results showed that the congruency effects (in terms of reduced saccade latencies for congruent versus incongruent trials) as observed in Experiments 3 and 5 (abstract geometric stimuli) do not vary as a function of a social versus non-social context.

### 3.2.3 DISCUSSION

Previous research on gaze interaction has mainly focused on cognitive processes in the gaze recipient, for example, by asking how humans respond to perceived gaze (gaze cueing, see Frischen et al., 2007, for a review). Here, I am addressing cognitive processes underlying gaze interaction from the actor's point of view and investigate how actors plan and initiate eye movements that trigger gaze responses in virtual faces of others. This novel approach has recently been proposed in the sociomotor framework (Kunde et al., 2018), an extension of the ideomotor view of action control (Harleß, 1861; Herbart, 1825; James, 1890; Lotze, 1852) to a social context, but has not been empirically addressed in the domain of gaze interaction, yet. Building on a pre-established paradigm for addressing anticipation-based motor control in other contexts, participants were asked to alternately look at two faces on the screen, which consistently responded to the participant's gaze with either direct or averted gaze (representing the gaze reaction of an interaction partner). Crucially, it was tested whether this gaze response of the targeted face is already anticipated prior to the participant's eye movement toward the target face. To do so, a task-irrelevant visual stimulus was displayed prior to the execution of the target saccade, which was either congruent, incongruent, or unrelated to the subsequently perceived gaze. Congruency effects on saccade latencies were interpreted as a marker of cognitive anticipation of the (virtual) other's gaze

response to the own gaze: Only when the subsequent effect of the participant's eye movement (toward one of the two faces) is anticipated, the congruency between the task-irrelevant visual stimulus and the anticipated effect has the potential to affect eye-movement control.

### **3.2.3.1 Acquisition of bi-directional action-effect associations**

The results clearly support my hypothesis: I observed a significant congruency effect on saccade latencies for schematic faces, which can be interpreted by assuming that participants indeed anticipated the gaze response of the targeted face prior to their target saccade toward that face. Thus, action-effect associations between the participant's gaze up/down and the subsequently perceived direct/averted gaze on the screen have been acquired and affected oculomotor control. These associations have not only affected saccade latencies, but also error rates as reflected in a significant effect of congruency on error rates in Experiment 5 (and a strong tendency toward a congruency effect in error rates in Experiment 3). As a consequence, I conclude that ideomotor control is not only at work in the guidance of eye movements in the inanimate environment (Huestegge & Kreutzfeldt, 2012) or in the context of emotion processing of the social environment (Herwig & Horstmann, 2011), but also plays an important role in gaze interaction, in line with the sociomotor framework (Kunde et al., 2018). Specifically, participants integrated the anticipated gaze response of their counterpart in oculomotor planning, which eventually controlled the perceived gaze response due to the eye movement-contingent display change technique.

While Experiment 3 involved schematic face stimuli with highly salient gaze responses, I did not find evidence of effect anticipation, neither in saccade latencies nor in error rates, when upright faces were used (Experiment 2). The lower saliency of the gaze response in photographic faces may account for the lack of more unambiguous effects in photographic faces (see below for a more elaborate discussion). While evidence for gaze cueing effects is usually found for a variety of different face/gaze stimuli (ranging from smiles and schematic faces to photographs of human and virtual avatar faces; see Lachat, Conty,

---

Hugueville, & George, 2012, and Frischen et al., 2007, for an overview), the presented data suggest that anticipating another person's gaze reaction is sensitive to the particular stimulus type used.

The lower saliency of the gaze in photographic faces may also account for the lack of congruency effects in Experiment 4, which involved inverted faces. Here, the inversion of the faces should disrupt holistic face processing (Bruce & Langton, 1994; Vecera & Johnson, 1995; Yin, 1969). This is in line with previous studies showing that gaze cueing effects were reduced or even absent (Kingstone, Friesen, & Gazzaniga, 2000; Langton & Bruce, 1999) when inverted faces served as stimuli (see Tipples, 2005, for a discussion of conditions under which orienting to eye gaze can be observed). In a similar vein, the study of Kunde et al. (2011) showed the absence of action-effect compatibility effects when emotional expressions of faces presented upside-down served as action effects.

I observed strong congruency effects in saccade latencies for the abstract geometric stimuli (Experiment 5). Here, visually simplified stimuli that resembled the gaze effects were used, but in a clearly non-social context. This effect conceptually replicates the findings of (Huestegge & Kreutzfeldt, 2012). However, I did not observe evidence of effect anticipation in Experiment 6 (scrambled face stimuli). This finding is of special interest, since the stimuli used in Experiment 6 differed from those of Experiment 3 (where I observed evidence of effect anticipation) only with respect to the configuration of object features while keeping overall visual low-level features constant. Probably, keeping low-level stimulus features such as color, contrast, pixel shifts etc. constant does not automatically ensure similar perceived conspicuity. Instead, it is possible that a small change of position of a group of pixels in a configuration interpreted as a face or eye region (schematic face, Experiment 3) is more easily perceived than a corresponding change in a random geometrical (meaningless) configuration (Experiment 6). Thus, only conspicuous effects may reliably establish a basis for ideomotor control mechanisms to come into effect.

Taken together, the results of Experiment 5 (abstract geometric stimuli) and Experiment 6 (scrambled face stimuli) allow the following conclusion to be drawn: While the action effect is a random (meaningless) visual change in both experiments, the stimulus materials of Experiment 5 (abstract geometric stimuli) and Experiment 6 (scrambled face stimuli) differed in one important aspect: The visual change is rather small (with respect to pixel alterations) in Experiment 6, but significantly larger in Experiment 5. It might be that this difference in size (and thereby in effect saliency) might be of utmost relevance for observing the congruency effect. Together with the present congruency effect in a gaze-interaction context (Experiment 3) and the absent evidence for effect anticipation in the scrambled face condition (Experiment 6), these data support the hypothesis that oculomotor action-effect associations are acquired in both types of environment (social vs. non-social) in a similar manner, albeit with some restrictions: If the visual action effect is rather unobtrusive, it must be integrated into a social context, thus increasing the subjective saliency of the visual action effect. If the visual action effect occurs in a meaningless non-social context, it must come with sufficient visual saliency with respect to low-level features, such as, for example, size, contrast, or spatial feature configuration, to affect saccade control. While I was only able to contrast a social context (Experiment 3) with a non-social context where the action effect was not meaningfully integrated into the surrounding visual features (Experiment 6), it would be interesting in the future to aim for an additional non-social control condition. Such stimuli would match the schematic face stimuli with respect to low-level properties while embedding the visual action effect into a more meaningful, but non-social context.

Note that Joyce et al. (2016) found evidence for anticipatory gaze cueing as reflected in manual RTs, which was driven by the mere prediction about the other person's gaze behavior. However, the study design and research question of Joyce et al. (2016) were different from the present approach in that they used an adapted standard gaze-cueing paradigm (with manual RTs as dependent measure) with a focus on the gaze recipient, while the present

---

design addresses the novel question of how humans control eye movements that trigger gaze responses in other people. Nevertheless, their main conclusion, namely that prior expectations regarding the gaze behavior of others can shape consecutive behavior, is consistent with what I have demonstrated for eye movements with the current study.

### 3.2.3.2 Temporal dynamics of effect anticipation in social versus non-social action contexts

Another aim of the present study was to assess the similarity or distinctiveness of the mechanisms driving anticipation-based control in the context of social and non-social action effects. Note that the following results relate to the between-experiment comparison of Experiments 3 and 5 since I observed strong congruency effects in these two experiments. The manipulation of the temporal delay between the visual stimulus and the imperative auditory stimulus (SOA) aimed at the investigation of potential differences in the *processing timeline*. Social action effects (e.g., a person reacting with a smile to another smiling person) tend to appear with a greater time delay than non-social action effects (e.g., a visual light flash as a result of a key press; see Kunde et al., 2018). Thus, I considered it plausible that anticipations of social action effects might need more time to emerge and tested whether the cognitive processing system may be attuned to such experiential differences in that congruency effects are more pronounced in the long versus short SOA condition, especially for the social context. However, while my results demonstrated that mean RTs were generally shorter in the long SOA condition compared to the short SOA condition for both social and non-social contexts, the congruency effect in RTs did not vary as a function of the SOA manipulation for social and non-social contexts, as revealed by the non-significant three-way interaction of experiment (Experiment 3 vs. Experiment 5), congruency, and SOA. Probably, 200 ms are already sufficient for a full build-up of anticipation, independent of a social versus non-social context. Specifically, mean baseline saccade latencies fell in between the latency levels of congruent and incongruent conditions for both SOA conditions. Consequently, I

conclude that the trial-based processing timeline appears to be highly comparable independent of the nature of the action effect (social vs. non-social). The absence of a difference in congruency effects depending on SOA and context could also indicate that the general idea of social action effects that tend to occur with a greater delay as compared to non-social action effects is neither very clear-cut nor strictly necessary. Social action effects can also be immediate, fast responses to another person's action, for example in case of the startle response (which represents a defensive response to threatening stimuli, such as when the other person performs a sharp movement close to my face). Conversely, it might sometimes also take some time until non-social action effects come into play, for example, the effect that turning on the heater has on the room temperature.

The main effect of SOA can be explained in the context of the typical time-course of temporal preparation (e.g., Bertelson, 1967; Bertelson & Tisseyre, 1969; Los & Schut, 2008) which has been shown to be comparable for both manual and oculomotor responses (Graefe & Vaughan, 1978). According to this research, the occurrence of a warning signal prior to an imperative stimulus facilitates RTs with an optimal task readiness at intervals of around 500 ms between the warning signal and imperative stimulus. Smaller RTs in the long SOA condition thus likely reflect better temporal preparation compared to the short SOA condition, which ultimately allows for faster saccade initiation in response to the imperative auditory stimulus (see also Kiesel & Hoffmann, 2004, for a similar result in the manual domain where responses were slowed down for short (100 ms) compared to long (1500 ms) task preparation intervals).

Note that the opposite temporal result pattern was observed in the study of Huestegge and Kreutzfeldt (2012), where the SOA = 400 ms condition produced prolonged RTs. This discrepancy might be attributed to a methodological difference: In the study by Huestegge and Kreutzfeldt (2012), the task-irrelevant stimulus was always presented for 400 ms, causing a temporal overlap between the visual presentation of the task-irrelevant stimulus



and the imperative auditory stimulus in the SOA = 200 ms condition, but not in the SOA = 400 ms condition. Therefore, the prolonged saccade RTs in the long SOA condition likely resulted from the additional processing time dedicated to the visual stimulus until the onset of the auditory stimulus. This additional processing may have interfered with rapid saccade initiation. In my work, by contrast, the presentation time of the task-irrelevant visual stimulus was decreased to 200 ms, resulting in a comparable presentation duration until the onset of the auditory stimulus. Note that the fixation cross reappeared in the SOA = 400 ms condition for an additional 200 ms between the offset of the visual stimulus and the onset of the auditory stimulus (see *Section 3.2.1* for further procedural details). This was done in order to avoid an additional temporal gap in the SOA = 400 ms condition where nothing except for the black screen would have been visible until fixation had reappeared with onset of the auditory stimulus, potentially causing a facilitation of saccadic RTs produced by stimulus offset (cf. Huestegge & Koch, 2010b; Kingstone & Klein, 1993; Saslow, 1967). Under the presented conditions, however, the preparation effect for SOA = 400 ms could likely be interpreted in an uncompromised manner.

### **3.2.3.3 Peculiarities and challenges for the investigation of gaze control in social contexts**

One aspect that is special for the social compared to the merely geometric action effects is that they can be valence-laden. For example, direct gaze can represent a (potentially rewarding) valence-laden approach signal. Moreover, eye contact can exert a powerful influence on cognitive processes like attention, evaluation, or memory (e.g., Böckler et al., 2014; Hood et al., 2003; for a review, see Senju & Johnson, 2009). The results of one study also suggested that action-effect learning could be modulated by the presence or absence of eye contact (Sato & Itakura, 2013, but see Riechelmann et al., 2019, for a failure to replicate this effect). In fact, direct gaze influenced saccade latencies in the sense that participants initiated their saccades faster when the task-irrelevant stimulus showed direct (vs. averted) gaze, but

this was also true for the respective non-social counterpart of the direct gaze face (the white ellipsis filled with a central black square). However, I could find neither an effect of the anticipated gaze type (direct vs. averted gaze) nor a combined influence of anticipated gaze type and congruency on RTs in Experiment 3 (schematic face stimuli). Together, these observations do not support the idea that eye contact substantially modulated the acquisition of action-effect associations in the present study.

Do these observations warrant the conclusion that sociomotor action control is inherently different from other, non-social forms of action control? In stimulus conditions in which clear congruency effects could be observed, namely with non-photographic stimuli, the presented data showed comparable effects in terms of temporal structure and effect size in a social and non-social context. Although one cannot perfectly compare the results from Experiments 3 (schematic face stimuli) and 5 (abstract geometric stimuli) due to low-level differences in the stimulus material, this observation speaks against the assumption of fundamentally separate mechanisms underlying socio- and ideomotor action control. The assumption of similar mechanisms for social and non-social processing contexts is further supported by results on reflexive orienting where a majority of studies comparing attentional cueing with gaze versus arrow stimuli suggest that both kinds of cues oriented attention in a comparable manner (Kuhn & Benson, 2007; Kuhn & Kingstone, 2009; Ristic, Friesen, & Kingstone, 2002; Tipples, 2002).

The absence of congruency effects for the photographic stimuli might be due to several aspects, which will be discussed more elaborately in the following. First, the gaze changes in these stimuli differed from the changes in the non-photographic stimuli with respect to low-level features. For example, the amount of pixels changing when shifting from the neutral baseline to a directional condition within the stimuli was lower for the photographic compared to the non-photographic stimuli. Thus, it is conceivable that in real-life, where faces of interaction partners are usually larger than the photographic stimuli on the

screen as used in the present experiments, gaze anticipation might play a more pronounced role than suggested by Experiment 2. Further, the black-and-white non-photographic stimuli were of higher contrast than the color photographic stimuli. High contrast is known to improve visual perception in general (e.g., Näsänen, Ojanpää, & Kojo, 2001), and, even more relevant in this case, to improve the perception of eye gaze direction when the area of the white sclera is larger (see Ricciardelli, Baylis, & Driver, 2000). Second, on a higher processing level more features can be processed in photographic stimuli compared to the non-photographic stimuli (e.g., gender, skin color, facial structure, and attractiveness), potentially claiming additional attentional resources that then are no longer available for gaze processing/anticipation. While I am aware that action-effect integration can principally occur in a relatively automatic manner (Elsner & Hommel, 2001; Hommel et al., 2003) this additional perceptual and attentional load might nevertheless have affected the results by strongly attenuating effects of congruency. All in all, I thus conclude that participants did not build up strong action-effect associations using photographic stimuli due to the lower overall saliency of the effect signal in these stimuli.

In addition, the claim that overall saliency of the gaze is reduced for photographic versus non-photographic stimuli empirically was addressed: In two short additional experiments, the detection rate of direct vs. averted (left/right) gaze for photographic (upright and inverted) face stimuli to the detection rate of direct vs. averted (left/right) gaze for schematic face and central vs. lateral effects for abstract stimuli using the stimulus material from the current study was determined. Participants either saw a stimulus presented for 35 ms at the screen center followed by a random pattern mask (scrambled version of the preceding stimulus) (Experiment A), or saw a printout of the stimuli (stimulus width  $\times$  height: 2.0 cm  $\times$  2.6 cm) at a fixed viewing distance that was large enough to prevent perfect stimulus classification (3.35 m, Experiment B). Participants had to indicate the (gaze) orientation of the presented stimulus (left vs. direct/central vs. right, using randomized stimulus placement). The

results confirmed the hypothesis that (gaze) orientation saliency is higher for non-photographic stimuli compared to photographic stimuli, as detection rate was higher for the abstract and schematic face stimuli compared to the photographic upright and inverted stimuli,  $t(19) = 6.85, p < .001, d = 1.53$  (in Experiment A), and  $t(19) = 6.51, p < .001, d = 1.46$  (in Experiment B).

### 3.2.3.4 Limitations

Finally, it should be noted that the presented results are interpreted within the framework of ideomotor control, even though they cannot be interpreted as direct evidence for a strong variant of ideomotor theory, which would posit that goal anticipation is *necessary* for action control (Shin et al., 2010). Specifically, the presented results only show that participants learned to associate certain pairings of effect stimuli and actions given that these stimulus-position pairings have been repeatedly experienced during the course of the experiment. In this way, association learning represents an underlying mechanism of ideomotor control. Crucially, however, in each trial the effect came *after* the action. Therefore, when assuming a very rudimentary learning mechanism, there would be no need for participants to represent the effect prior to action execution mentally, as this is not necessary to produce the required action in the first place. Thus, the results cannot simply be explained with a more basic associative learning mechanism, as they clearly show that effect representation is already present prior to the action. This conclusion is essentially the tenet of what has been referred to as the “weak” variant of ideomotor theory (cf. Huestegge & Kreutzfeldt, 2012).

In conclusion, the presented findings demonstrate that humans integrate the social consequences of oculomotor actions (in terms of the anticipated visual gaze response of a virtual interaction partner), which can, in turn, support shaping other humans’ gaze behavior by one’s own oculomotor actions. The results can be embedded into the sociomotor framework of human action control (Kunde et al., 2018), which discusses potential peculiarities of human action control in a social context, *inter alia*, with respect to the role of eye movements

---

as a specialized effector system. Accordingly, the findings also indicate that effect-based oculomotor control is similar in social and non-social action contexts.

### 3.3 CONCLUSION OF CHAPTER 3

In general, the experiments presented in this chapter investigated whether learned associations between oculomotor actions and their visual effects determine saccade control in both non-social (Experiments 1 and 4–6) and social environments (Experiments 2 and 3). Albeit this common goal in mind, I used two different experimental approaches to address the overarching research question. While Experiment 1 focused more closely on the spatio-temporal dynamics of the mechanisms underlying saccade control in a non-social context, the goal of Experiments 2–6 was a direct comparison of the nature of ideomotor and socio-motor saccade-effect associations within the identical experimental paradigm. Such a direct contrast comes with the great advantage of a high degree of comparability, since manipulations are specifically targeted at the variable of interest, while the remaining factors are kept constant. Moreover, Experiments 2 and 3 served to address gaze interaction from the viewpoint of the agent – an aspect of sociomotor gaze control that has not yet been investigated. Therefore, participants in all experiments were required to perform saccades toward certain saccade targets. Specific and contingent visual effects in either objects (non-social context) or faces (social context) followed these saccades. If saccades and their effects can become associated through repeated experience of such saccade-effect associations in a bi-directional manner, as suggested by ideomotor theory and the sociomotor framework of action control, this link should be reflected in certain saccade parameters. The critical question of this chapter was whether the nature of these oculomotor action-effect associations is similar for social vs. non-social environments.

Taking everything into account, the results of the current chapter clearly support the hypothesis that effect-based oculomotor control is similar across social and non-social contexts. More precisely, the results of Experiment 1 indicated that the anticipation of oculomotor action effects is reflected in spatial saccade characteristics in inanimate environments, similar to corresponding effects demonstrated in a social context (Herwig & Horstmann, 2011). The combined results of Experiments 2–6 demonstrated congruency effects (as an indicator of the acquisition of bi-directional associations between saccades and their effects) for both social and non-social stimuli, but only when the perceived changes are sufficiently salient. Temporal dynamics of the congruency effects were comparable for social and non-social stimuli, suggesting that similar mechanisms underlie effect-based oculomotor control. Overall, these results support recent theoretical claims emphasizing that similar ideomotor processes as involved to affect the inanimate environment in desired ways are also employed when it comes to action control in a social context (Kunde et al., 2018).

A methodological concern might relate to the stimulus material used in this chapter to establish a social or non-social context, respectively. One might argue that presenting faces on the screen does not really establish a social situation comparable to the presence of real human interaction partners. Moreover, the presence (vs. absence) of real persons has the power to shape human behavior, for example with respect to helping behavior (also referred to as the bystander effect, see Hortensius & Gelder, 2018, for a recent review), but also when looking at rather basic human behavior like attentional mechanisms (Wühr & Huestegge, 2010). Consequently, the question arises whether the present experimental context allows drawing conclusions about potential peculiarities of social action control. Since this methodological issue is relevant throughout this entire work, I will further elaborate on this point in the General discussion (*Section 7.3.1*).

According to the sociomotor framework (Kunde et al., 2018), social action effects are usually less contingent upon an agent's action and typically occur with a greater delay

---

than non-social action effects (see *Section 2.1.3*). However, the action effects as studied in this chapter occurred perfectly predictable upon the executed action. While the action-effect interval was manipulated in Experiment 1, where I exclusively relied on non-social stimuli as action effects, it was kept constant across Experiments 2–6, where I directly contrasted ideomotor and sociomotor gaze control. Thus, pending issues that need to be addressed with respect to commonalities and differences in ideomotor and sociomotor gaze control include the investigation of the role of effect contingency and effect contiguity in social vs. non-social action control contexts. Within the subsequent chapter of this work, I will try to tackle this research question.

## 4 DETERMINANTS OF ACTION-EFFECT LEARNING IN OCULOMOTOR ACTION CONTROL

### 4.1 EFFECT CONTINGENCY AND EFFECT CONTIGUITY IN SOCIAL VERSUS NON-SOCIAL CONTEXTS

As outlined in *Section 2.2*, existing research on ideomotor action control has already focused on the role of both effect contingency (e.g., Elsner & Hommel, 2004) and effect contiguity (e.g., Dignath et al., 2014; Dignath & Janczyk, 2017; Elsner & Hommel, 2004, for the manual domain; see *Section 3.1* of the present work for the oculomotor domain) on the acquisition and bi-directional retrieval of action-effect associations. These findings allow us to draw the general conclusion that action-effect learning seems to benefit from both high contingency and high contiguity (Elsner & Hommel, 2004). Note that these findings were obtained from studies conducted in inanimate environments where participants experienced that their actions (e.g., a key press) triggered certain perceptual effects (e.g., tones). These results are informative given that humans perform such kind of actions several times a day in their daily life, for example, when pressing the power button to turn on the TV or pressing the light switch to turn on the light. However, as described in the sociomotor framework of action control (Kunde et al., 2018), humans do not exclusively interact with inanimate things, but are rather surrounded by other people, and their behavior inevitably affects the behavior of these others (see *Section 2.1.3* for a more detailed elaboration on the nature of social action effects). Since social action effects come with some peculiarities when it comes to effect



---

contingency and effect contiguity, it seems reasonable to investigate whether the role of action-effect contingency and action-effect contiguity differs in social vs. non-social environments. In the following paragraph, I will elaborate on the aforementioned peculiarities of social action effects with respect to effect contingency and effect contiguity and their implications for action control.

Even though there are quite a lot of situations where the behavior of one person elicits a (perfectly) predictable response in another person, one can easily think of counterexamples. For example, waving to someone else might very likely cause the other person to respond with some kind of positive feedback indicating that he or she has noticed the waving person, but this behavioral response can come as a positive expression other than waving, like saying hello, nodding of the head, or smiling at the waving person. However, it might also happen that the other person does not respond at all or responds in a rather unfriendly way for some reasons not apparent to the waving person (e.g., he or she might have had a bad day at work or felt sick that day). Thus, social action effects generally tend to be less predictable, that is less contingent on our own actions compared to most non-social action effects. The nature of social and non-social action effects also differs with respect to the evaluation of potential outcomes: While the actor in general accepts that there is a bunch of individual response possibilities toward her/his own action in social interactions, there is much less (or even no) tolerance when it comes to action effects in the non-social environment. Consider the light switch example from above: The two potential effects of someone pressing the light switch could be that the lights go on or remain dark, but only the former effect is accepted by the actor. The typical knocking sound when knocking on a table is an even more fundamental example which shows that most actions in the inanimate environment come with a very specific expectation about the sensory consequences of the respective action. Hence, Kunde and colleagues (2018) argue that the level of contingency might affect action-effect learning differently depending on whether the action effect results from a social

interaction partner or from the inanimate environment. More precisely, they hypothesize that the generalization across different types of individual action effects might be stronger depending on whether these effects are interpreted as originating from a social rather than a non-social source (see Schilbach, 2014, for a similar discussion).

While non-social action effects are typically the direct, mechanical result of an actor's own action (e.g., the sound that results from pressing a certain piano key), social action effects are executed by another person. Therefore, it is reasonable to assume that social action effects tend to come into effect with reduced temporal contiguity as compared to action effects in the inanimate environment (Kunde et al., 2018). Using the examples from above, I will explain the underlying rationale more clearly: The auditory effect of knocking on a table occurs instantaneously after doing so. However, with respect to the waving example, it takes some time for the person who was waved at to register and process the waving in order to respond accordingly. Given this inherent temporal difference in the occurrence of social vs. non-social action effects, the question arises whether the influence of the temporal delay in social environments differs from what has been reported for the influence of the temporal delay in inanimate environments. While high temporal contiguity has shown to facilitate the acquisition of associations between motor actions and non-social effects in general (Elsner & Hommel, 2004, but see Dignath et al., 2014), it might be that longer delays between action and effects are more optimal to acquire action-effect associations for social as compared to non-social effects.

In a similar vein, research on imitation, which can be defined as the behavioral tendency to copy observed actions (Heyes, 2011), has also focused on the behavioral effects of contingency and contiguity, since imitation has been shown to also result from associative, experience-based learning mechanisms (Heyes, 2001; Heyes, Bird, Johnson, & Haggard, 2005), similar to the acquisition of action-effect associations. Thus, several imitation studies tested whether the basic learning principles of associative learning – that is contingency and

---

contiguity – affect the effects of imitation (Bailenson, Beall, Loomis, Blascovich, & Turk, 2004; Catmur & Heyes, 2013; Dignath, Lotze-Hermes, Farmer, & Pfister, 2018). The study of Dignath et al. (2018), for example, provided evidence that contingency and contiguity affect the social effects of imitative behavior in terms of altered social affiliation judgments toward the interaction partner. When contingency and contiguity between the action executed by the participant and the imitative response executed by the model were high, participants judged the model as more likeable as compared to when contingency and contiguity were low. These findings are of interest for the present work, since they bridge research on sociomotor action control with imitation research by demonstrating that basic learning mechanisms – contingency and contiguity – mediate social effects of imitation in terms of altered social affiliation in the actor.

These considerations were the starting point for the following series of experiments aiming at a systematic comparison of the impact of contingency and contiguity on oculomotor action-effect learning with respect to context-specific effects (social vs. non-social). The procedure of the present experiments was based on the procedure of Experiments 2–6, albeit with some changes. First, the SOA manipulation was discarded based on the results of Experiments 2–6, where congruency effects emerged independent of the SOA conditions. Thus, the SOA was kept constant at 400 ms in Experiments 7 and 8. Second, I eliminated the factor *direction* of the action effect from analysis (direct vs. averted gaze for social stimuli; central vs. lateral effect for non-social stimuli) in order to focus the analysis on the most relevant factors and to avoid false-positive results because of multiple testing.

## 4.2 EXPERIMENT 7A

Experiment 7a was set up to replicate the congruency effect as observed in Experiments 3 and 5 for schematic face and abstract stimuli, while investigating the role of (perfect versus reduced) contingency on the acquisition of action-effect associations. Experiment 7a also asked whether contingency affected action-effect learning differently across a social versus a non-social context. Therefore, I used a similar setting as in Experiments 2–6, but manipulated the level of contingency between oculomotor actions and their effects. While half of the participants was treated in exactly the same way as in Experiments 2–6 in that they experienced perfect contingency between their eye movements and the visual effects (e.g., every saccade to one face location was followed by perceiving direct gaze), contingency was reduced in the remaining half of participants (e.g., a saccade to one face location resulted in the more probable effect in only 80% of all trials).

Elsner and Hommel (2004) showed learning of action-effect associations to be impaired when the co-occurrence of action and effect is reduced<sup>11</sup>. Thus, I expected congruency effects to be larger for the perfect versus reduced contingency group. Given that social interactions are generally characterized by a reduced level of contingency (e.g., smiling at someone might trigger a smile in the other person, but also any positive response other than smiling, or no response at all), one might expect that the impact of reduced contingency is attenuated for social (vs. non-social) action effects.

---

<sup>11</sup> In the study of Elsner and Hommel (2004), the frequency and probability of the presence or absence of an effect tone given the presence or absence of a key press in the acquisition phase was manipulated. In different contingency groups, contingency was lowered by increasing the probability of an effect tone given the absence of a key press while keeping constant the probability of an effect tone given the occurrence of a key press. In further contingency groups, the effect tone appeared equally often in the presence and in the absence of a key press, but the overall frequency of the effect tone was varied such that the effect tone was presented in 80/50/20% of the trials during acquisition. Results indicated that action-effect learning was pronounced if the effect tone only rarely appeared in the absence of a keypress and if the overall frequency of the effect tone was high.

## 4.2.1 METHODS

### 4.2.1.1 Participants

Based on the reasoning from Experiments 2–6, where a power analysis (Faul, Erdfelder, Lang, & Buchner, 2007) indicated a sample size of six participant as sufficient to observe a congruency effect with considerable probability (power = .95,  $\alpha$  = .05), I tested 12 participants per each group, resulting in a total sample size of  $N = 48$  participants. Thus, I recruited 48 participants (mean age = 23.0 years,  $SD = 4.4$  years, age range: 18 – 35 years, 9 male). All participants reported normal or corrected-to normal vision and were naïve with respect to the purpose of the experiment. Prior to the experiment, all participants gave informed consent. They received either course credits or monetary compensation for participation.

### 4.2.1.2 Stimuli and procedure

For the selection of the stimulus material, the results of Experiments 2–6 were taken into account. Hence, I used the schematic face stimuli with direct gaze, averted gaze, and closed eyes (baseline) for the social context, and the abstract elliptical stimuli with a central back square, a lateral black square and no black square (baseline) for the non-social context, since they turned out to be the most optimal stimuli to observe congruency effects in Experiments 2–6 (see *Figure 8*). Stimuli were of the same size as in Experiments 2–6. The experimental design was highly similar to the design of Experiments 2–6, but differed in terms of the SOA and the contingency manipulation. In contrast to Experiments 2–6, the duration of the temporal interval between the task-irrelevant visual and the auditory stimulus (i.e., SOA) was kept constant at a duration of 400 ms. To manipulate contingency, participants were assigned to either the perfect or the reduced contingency group. For the perfect contingency group (as in Experiments 2–6), all saccades to one target location were followed by perceiving the same effect (see *Figure 11*). For the other group (reduced contingency), a saccade to one target location resulted in the more probable effect in 80% of all trials, while the

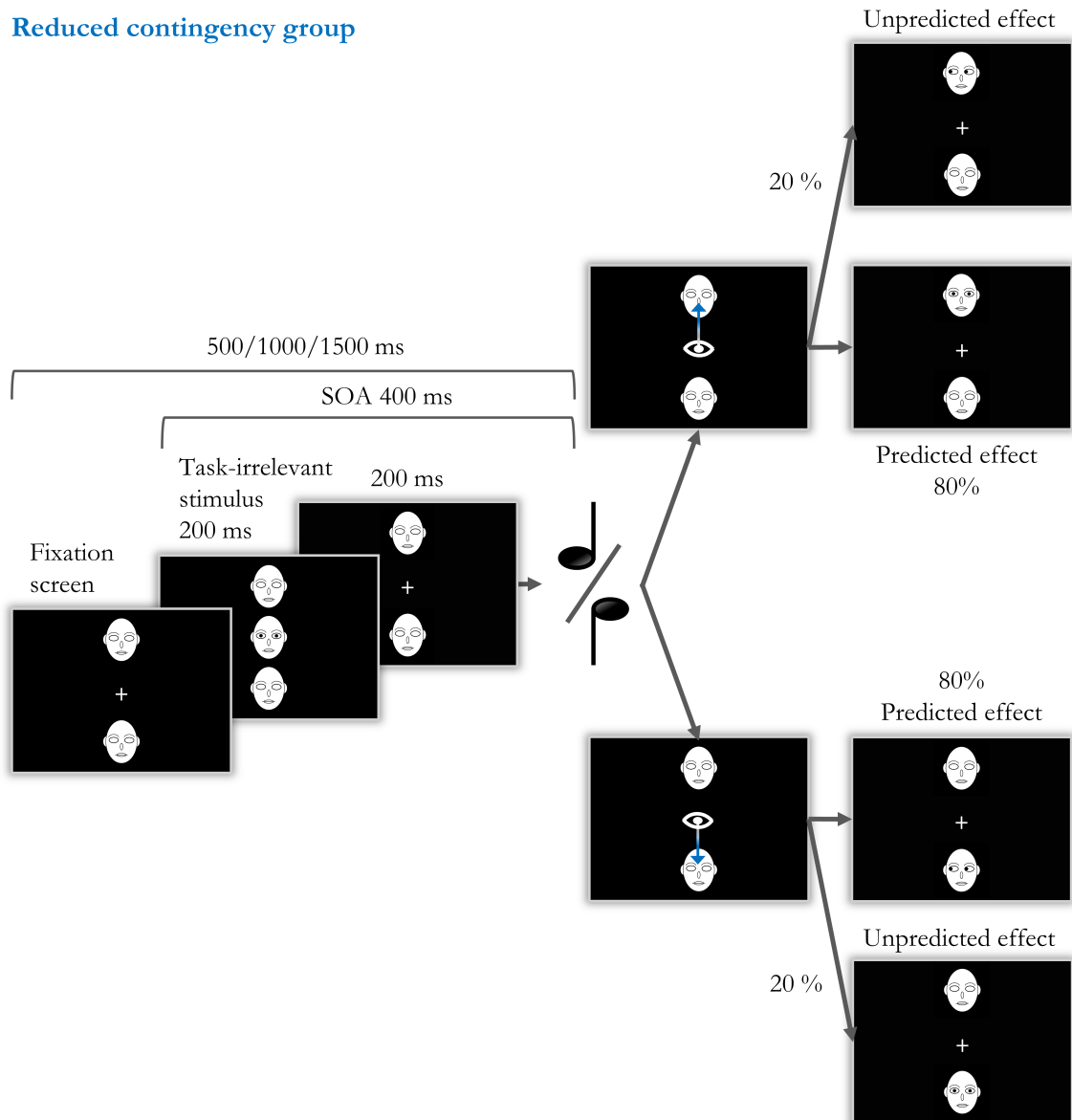
effect associated with the other target location occurred in the remaining 20% of trials. This means that, for example, the on-screen face responded with an averted gaze in one target location in 80% of trials, and the same face responded with direct gaze in the remaining 20% of trials toward this target location. A post-survey at the end of the experiment tested whether participants consciously noticed the contingency between the direction of their eye movement and the change within the saccade target. The post-survey followed a three-step procedure, which was stopped at a certain step, once the participant was able to recall the experienced action-effect contingency. First, participants were asked what they had noticed during the experiment. In a second step, this is, in case they did not report the correct action-effect contingency in the previous step, they were asked whether they had noticed any regularities between their eye movements and the on-screen changes. If the answer was “yes”, they were asked to recall what they had noticed. Finally, if the previous answer did not correspond with the actual action-effect contingency, or if the participants’ answer to the second question was “no”, they were directly asked whether they noticed that there was a (perfect/reduced) contingency relationship between their eye movements and the resulting effects.<sup>12</sup>

The experiment (around 35 min.) comprised 15 experimental blocks with 30 trials each (450 trials in total). Prior to each block, a calibration of the eye tracker was performed. The trial order was pseudo-randomized to prevent order effects. Thus, the amount of changes between high-pitched and low-pitched imperative auditory stimuli was equally distributed. The maximum sequence length of the same-pitched auditory stimulus was restricted to seven consecutive trials.

---

<sup>12</sup> It should be noted that the original question contained a detailed description of the experienced (perfect vs. reduced) contingency without using the technical term of “contingency”, but was simplified in the main text to increase readability.

### Reduced contingency group



*Figure 11.* Trial structure in Experiments 7a and 7b depicted for the reduced contingency group with schematic face stimuli. After the presentation of a black screen with a white fixation cross and two peripheral baseline face stimuli (without gaze), the fixation cross was replaced by a task-irrelevant stimulus, which was either a baseline face (without any visible gaze) stimulus or a face with direct or averted gaze. After that, a tone (high vs. low) signaled the participant to respond with a corresponding gaze shift (up vs. down). As an effect of the saccade toward the upper or lower face, either a direct or an averted gaze was displayed. In the “perfect contingency group”, each saccade triggered the onset of an action effect, which was perfectly contingent upon the targeted face. In the crucial “reduced contingency group” (as depicted in the present figure), the more probable action effect appeared in 80% of the trials (e.g., direct gaze when the upper face was targeted), while the less probable action effect (e.g., averted gaze when the upper face was targeted) was displayed in the remaining 20% of trials. In contingent trials, the central visual stimulus was either congruent, incongruent, or unrelated to the subsequent gaze response of the targeted face. For contingent trials, the trial sequence shown in this figure depicts a congruent and incongruent trial sequence with a direct gaze as the task-irrelevant stimulus. In case of an averted gaze as task-irrelevant stimulus (not depicted in the present figure), a gaze shift toward the upper target face (triggering direct gaze) would represent an incongruent trial, and a gaze shift toward the lower target face (triggering averted gaze) would represent a congruent trial. For the unrelated congruency condition, the task-irrelevant stimulus was a baseline face without any visible gaze.

### 4.2.1.3 Design and analysis

Repeated-measures ANOVAs with the within-subjects factor *congruency* (congruent vs. incongruent vs. unrelated) and the two group factors *contingency* (perfect vs. reduced) and *context* (social vs. non-social context) were conducted for saccade latencies and oculomotor error rates.

## 4.2.2 RESULTS

Prior to analysis, I excluded all non-contingent trials of the reduced contingency group (equivalent to 90 trials out of 450 trials per participant) which left 19,440 trials to be analyzed. I applied the same cleansing procedure as described for Experiments 2–6 (see *Section 3.2.2*), which led to the exclusion of 528 trials in total, corresponding to 2.7% of 19,440 trials.

### 4.2.2.1 Temporal oculomotor performance

*Figure 12* shows participants' saccade latencies in congruent, incongruent and unrelated trials separately for groups with perfect and reduced contingency. The repeated-measures ANOVAs with the factors congruency, contingency, and context revealed a significant effect of congruency,  $F(2, 88) = 10.96, p < .001, \eta^2_p = .20$ , indicating reduced saccade latencies for the congruent condition ( $M = 243$  ms,  $SE = 6$  ms) compared to both the incongruent condition ( $M = 251$  ms,  $SE = 6$  ms;  $p < .001$  for the pairwise comparison) and the unrelated condition ( $M = 248$  ms,  $SE = 6$  ms;  $p = .010$ ). Saccade latencies were also significantly reduced for the unrelated condition compared to the incongruent condition,  $p = .027$ . Note that none of the remaining statistical comparisons reached statistical significance, all  $F_s < 1$  (except for the three-way interaction with  $F(2, 88) = 1.17, p = .315, \eta^2_p = .03$ ). Thus, there was no significant main effect or interaction related to the critical contingency manipulation.



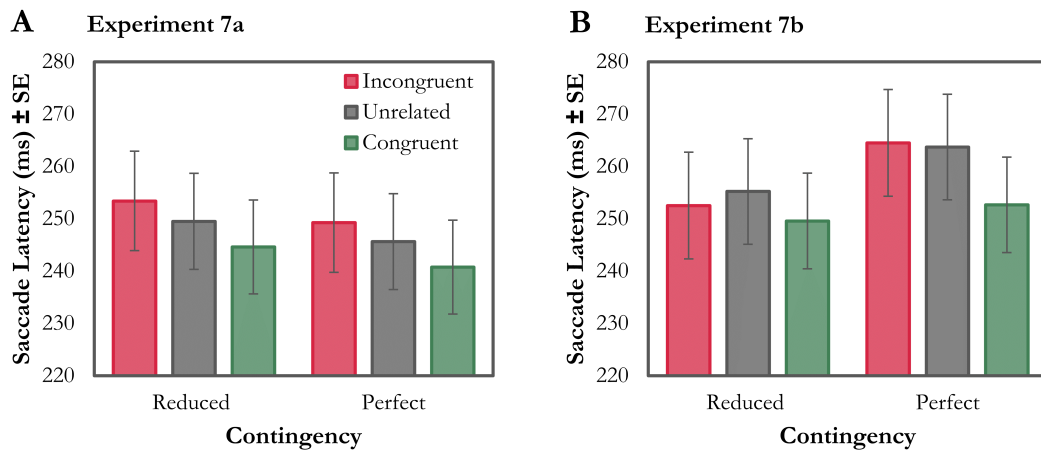


Figure 12. Saccade latencies (in ms) as a function of congruency (congruent vs. incongruent vs. unrelated) and contingency (reduced vs. perfect) (A) in Experiment 7a, and (B) in Experiment 7b, averaged across context condition. Error bars represent standard errors of the mean ( $SE$  in ms).

#### 4.2.2.2 Oculomotor error rates

On average, participants committed 6.7% ( $SE = 0.8\%$ ) errors in the congruent condition, 8.7% ( $SE = 1.0\%$ ) errors in the incongruent condition, and 8.3% ( $SE = 0.9\%$ ) in the unrelated condition, yielding a significant main effect of congruency,  $F(2, 88) = 9.36$ ,  $\varepsilon = .86$ ,  $p < .001$ ,  $\eta^2_p = .18$ . The error rate was significantly reduced for the congruent condition compared to both the incongruent condition,  $p = .001$ , and the unrelated condition,  $p < .001$ , while the error rate did not differ between the incongruent and the unrelated condition,  $p = .366$ . Neither the interaction of contingency and congruency was significant,  $F(2, 88) = 2.66$ ,  $\varepsilon = .86$ ,  $p = .084$ ,  $\eta^2_p = .08$ , nor any of the remaining statistical comparisons, all  $F$ s  $< 1$  (except for the three-way interaction with  $F(2, 88) = 1.39$ ,  $p = .256$ ,  $\eta^2_p = .03$ ). Thus, again there was no evidence for a significant impact of contingency in the data.

#### 4.2.2.3 Exploratory analyses

To gain further insight into the congruency effect and potential interactions with different levels of contingency, I conducted several exploratory analyses with respect to temporal aspects of the congruency effect. Firstly, it was explored whether an influence of action-effect contingency on the congruency effect separately for fast and slow RTs exists (*speed*

---

*analysis*). Therefore, I categorized saccade RTs as high or low using a median split analysis separately for each participant and condition. RTs and error rates were subsequently submitted to four-way repeated-measures ANOVAs with speed (fast vs. slow) as an additional within-subjects factor. Note that I will only report effects involving the relevant factor speed here. In order to concentrate on the relevant difference between the congruent and incongruent condition and to simplify analyses, I only implemented these two conditions into the following analyses, while omitting the data of the unrelated congruency condition. Consequently, congruency effects refer to reduced saccade latencies for congruent versus incongruent conditions.<sup>13</sup> Secondly, I explored the temporal dynamics of the congruency effect in two ways to assess whether the congruency effect develops faster for the perfect versus reduced contingency group, and whether this pattern was different across the social versus non-social context conditions. On the one hand, I conducted separate repeated-measures ANOVAs with congruency, contingency, and context for RTs and error rates involving the initial 30 trials only (*initial performance analysis*). On the other hand, I conducted ANOVAs involving the additional factor block (1–5; with each block involving 90 trials) to explore the temporal dynamics underlying the congruency effect (*temporal dynamics analysis*). For the sake of brevity, I only report effects involving the relevant factor block for the block analysis, as the result pattern for non-block related effects is identical with what has been reported above in the context of the three-way ANOVAs.

#### *Speed analysis*

Analyzing saccade latencies, I observed a (self-evident) effect of speed,  $F(1, 44) = 368.94, p < .001, \eta^2_p = .90$ . The interaction of congruency and speed,  $F(1, 44) = 3.15, p = .083, \eta^2_p = .07$ , as well as the four-way interaction,  $F(1, 44) = 3.87, p = .056, \eta^2_p = .08$ , were

---

<sup>13</sup> Please note that analyses including the unrelated condition overall yielded the same result patterns as the reduced analyses (i.e., without the unrelated congruency condition).

not significant. No other statistical comparison involving the factor speed reached significance, all  $F$ s  $< 1$  (except for the interaction of congruency, context and speed,  $F(1, 44) = 1.23$ ,  $p = .273$ ,  $\eta^2_p = .03$ , and of congruency, contingency and speed,  $F(1, 44) = 2.08$ ,  $p = .156$ ,  $\eta^2_p = .05$ ).

Regarding error rates, a significant effect of speed,  $F(1, 44) = 10.44$ ,  $p = .002$ ,  $\eta^2_p = .20$ , demonstrated that participants committed more errors at fast RTs ( $M = 9.9\%$ ,  $SE = 1.2\%$ ) compared to slow RTs ( $M = 5.5\%$ ,  $SE = 1.0\%$ ). Further, the interaction of congruency and speed was significant,  $F(1, 44) = 6.95$ ,  $p = .012$ ,  $\eta^2_p = .14$ . While participants showed a clear congruency effect at faster RTs in terms of significantly fewer error rates for congruent vs. incongruent trials,  $p = .001$ , the congruency effect was absent at slower RTs,  $p = .320$ . No other effect involving the factor speed was significant, all  $F$ s  $< 1$ .

#### *Initial performance analysis*

For the analysis of initial RTs, one participant had to be excluded due to empty cells in one of the congruency conditions. When analyzing initial RTs, none of the effects reached significance: Neither the effect of congruency,  $F(1, 43) = 3.15$ ,  $p = .083$ ,  $\eta^2_p = .07$ , nor the interaction of contingency and context,  $F(1, 43) = 3.93$ ,  $p = .054$ ,  $\eta^2_p = .08$ , were significant. Also, none of the remaining statistical comparisons reached significance, all  $F$ s  $< 1$  (except for the effect of contingency,  $F(1, 43) = 1.90$ ,  $p = .175$ ,  $\eta^2_p = .04$ ).

The analysis of initial error rates showed that the congruency manipulation did not affect error rates,  $F < 1$ . There was no significant interaction of congruency and contingency,  $F(1, 43) = 2.28$ ,  $p = .139$ ,  $\eta^2_p = .05$ , and no significant effect of contingency,  $F(1, 43) = 1.83$ ,  $p = .183$ ,  $\eta^2_p = .04$ , and all other effects returned non-significant results, all  $F$ s  $< 1$ .

#### *Temporal dynamics analysis*

Analyzing saccade latencies block-wise using a four-way repeated-measures ANOVA with block as additional factor revealed a significant main effect of block,  $F(4, 176) = 9.59, \varepsilon = .75, p < .001, \eta^2_p = .18$ . Pairwise comparisons showed that RTs tended to decrease over the course of the experiment ( $ps < .05$  for the pairwise comparisons between blocks, except for the comparison of Block 1 and 2,  $p = .052$ , Block 2 and 3,  $p = .146$ , and Block 4 and 5,  $p = .670$ ). Neither the interaction of congruency and block,  $F(8, 176) = 1.09, \varepsilon = .77, p = .358, \eta^2_p = .02$ , of contingency and block,  $F(4, 176) = 1.22, \varepsilon = .75, p = .306, \eta^2_p = .03$ , nor the three-way interaction of block, contingency and context,  $F(4, 176) = 1.25, \varepsilon = .75, p = .295, \eta^2_p = .03$ , nor any other interaction involving the factor block reached significance, all  $F_s < 1$ .

When submitting error rates to a four-way ANOVA with block as additional factor, no effect involving the factor block was significant (all  $F_s < 1$ , except for the interaction of block and contingency,  $F(4, 176) = 1.17, \varepsilon = .61, p = .319, \eta^2_p = .03$ , the three-way interaction of block, contingency and context,  $F(1, 176) = 1.04, \varepsilon = .61, p = .368, \eta^2_p = .02$ , and the three-way interaction of block, congruency and context,  $F(1, 176) = 1.06, \varepsilon = .84, p = .370, \eta^2_p = .02$ ).

### 4.2.3 DISCUSSION

The purpose of the present experiment was to investigate whether the influence of reduced action-effect contingency on action-effect learning is attenuated for social versus non-social action effects, as suggested by the sociomotor view on action control. This theoretical prediction was based on the observation that social action effects tend to come with reduced contingency compared to most non-social perceptual action effects. However, the results showed that both saccade latencies and error rates were unaffected by the contingency manipulation across both contexts, as revealed by the non-significant effect of contingency and the lack of significant interactions with contingency. This means that the strength of the

---

action-effect associations acquired through repeated experience of the link between performing an action and observing its perceptual effect was (statistically) indistinguishable for both contingency groups. This leads me to reject the hypothesis of an attenuated influence of contingency for a social versus non-social context.

However, the results of Experiment 7a demonstrated a replication of the context-independent congruency effect as observed in Experiments 3 and 5, such that participants were faster (and less error-prone) to initiate a saccade toward a target when a task-irrelevant stimulus, presented prior to the execution of the target saccade, was congruent rather than incongruent to the subsequently perceived target. These results are in line with the basic ideomotor assumptions that participants acquired bi-directional action-effect associations, and that the presentation of the action effect in terms of a task-irrelevant stimulus prior to saccade execution triggered the associated action.

The results of the exploratory analyses revealed some interesting insights into the nature of the present findings. First, none of these analyses showed any influence of the contingency manipulation on the observed results. Second, the speed analysis demonstrated a speed-accuracy trade-off, since participants committed more errors at low (versus high) RTs. Moreover, the congruency effect in error rates was modulated by response speed: A clear congruency effect was evident in trials where participants reacted rather fast, while such an effect was absent in trials with slow RTs. This finding points toward a rather reflexive and automatic character of the response tendency as activated by the task-irrelevant prime. If participants take more time to execute the response saccade, they are able to suppress the prime-associated response in case of an incongruent trial. Interestingly, this suppression was not reflected in saccade latencies, where congruency effects tended to occur independent of the speed of response. Further, congruency effects in both saccade latencies and error rates were not present after the first few trials, suggesting that a repeated experience of the action-effect associations is necessary to elicit an activation of the action effect via the task-irrelevant

prime. Again, this finding was independent of the contingency or context manipulation. This observation contradicts the results of studies that demonstrated a fast learning of such action-effect associations (Hommel, 2009; Pfeuffer et al., 2016), but the reasons for this could be located in the particular design of the present task. While experimental designs addressing the acquisition of action-effect associations usually consist of two distinct phases, that is, an acquisition and a test phase (see *Section 2.1.2.1* for a detailed description of the action-effect acquisition paradigm), the present design interweaves these two phases into a one-phase design, where the stimulus supposed to prime the associated response is presented from the very first trial on. This additional visual (and task-irrelevant) input might hinder the fast build-up of action-effect associations as reflected in congruency effects.

The analysis of the temporal dynamics of the congruency effect only demonstrated decreasing RTs over the course of the experiment. This indicates that participants became more and more familiar with the task.

At that point, it is interesting to report that the majority of participants of both contingency groups in Experiment 7a were not able to recall the experienced contingency relationship after the experiment. In the post-survey, only 12 out of 48 subjects (five of the reduced and seven of the perfect contingency group, respectively) reported the correct action-effect relationship when they were asked what they noticed during the experiment. Another nine subjects (six of the reduced and three of the perfect contingency group, respectively) reported the correct action-effect relationship when they were asked about any regularities between their eye movements and the on-screen changes. Of those who were not able to recall the experienced contingency relationship in the first two steps of the post-survey (corresponding to 27 participants), 15 subjects (seven of the reduced and eight of the perfect contingency group, respectively) responded with “no” when directly asked whether they noticed that there was a (perfect/reduced) contingency relationship between their eye

---

movement and the resulting effect. These numbers show that action-effect learning, as reflected in stable congruency effects, had mainly occurred implicitly, as a majority of participants did not consciously capture the relation between performing an eye movement toward a certain target and the visual effect resulting from this action, independent of the contingency group. The fact that the action effects were of no task relevance for the participants might have contributed to this phenomenon.

Taken together, the results of Experiment 7a demonstrated a strong congruency effect as an indicator of the context-independent acquisition of action-effect associations. However, action-effect contingency, the main variable of interest in this experiment, showed no influence on the congruency effect, such that the research question that social versus non-social environments deal differently with reduced contingency could not be answered. Potential reasons for the absence of any influence of action-effect contingency will be discussed in more detail within the discussion section of Experiment 7b (*Section 4.3.3*).

## 4.3 EXPERIMENT 7B

Experiment 7b was conducted in parallel with Experiment 7a and the procedure of both experiments was identical except the following alteration: While learning of the contingencies between actions and effects was implicit in Experiment 7a, participants were explicitly instructed about the contingency relations in Experiment 7b. This explicit emphasis on the action-effect contingencies was intended to increase action-effect learning. Additionally, I emphasized the participant's role as the cause of the action effects in both contingency conditions (e.g., "Throughout the experiment, your eye movements will always elicit an effect in the upper or lower face/object, depending on the direction of your eye movement"). Such kind of instructions are known to increase the impact of action effects on action control (Janczyk, Yamaguchi, Proctor, & Pfister, 2015). A comparison between Experiments 7a and 7b allowed assessing the extent to which explicit instruction is beneficial for the learning of action-effect associations in conjunction with perfect versus reduced action-effect contingency. Hence, I expected contingency-related differences in congruency effects to be reduced or even absent when participants are explicitly instructed about the action-effect contingencies (Experiment 7b) versus when learning of the action-effect contingencies occurs implicitly (Experiment 7a).

### 4.3.1 METHODS

#### 4.3.1.1 Participants

Forty-eight participants were tested (mean age = 23.5 years,  $SD = 3.9$  years, age range: 18 – 33 years, 9 male) in Experiment 7b. Every participant reported normal or corrected-to normal vision and was naïve with respect to the purpose of the experiment. All participants gave informed consent before the experiment started, and received either course credits or monetary compensation for participation.



#### 4.3.1.2 Stimuli and procedure

Stimuli, procedure, design, and analysis were identical to Experiment 7a except one modification: In contrast to Experiment 7a, participants were explicitly instructed about the perfect or reduced contingency condition in Experiment 7b. These instructions were presented on the screen just before the experiment started, and participants were asked by the experimenter to verbally repeat the contingency relation in order to ensure that they understood the contingency relations. In case they could not correctly reproduce the contingencies, the experimenter explained the contingency relation once again.

#### 4.3.1.3 Design and analysis

I conducted a between-experiment comparison including the within-subjects factor *congruency* (congruent vs. incongruent vs. unrelated) and the group factors *contingency* (perfect vs. reduced), *context* (social vs. non-social context), and *experiment* (Experiment 7a and 7b) within a mixed ANOVA to assess the extent to which explicit instruction is beneficial for the learning of action-effect associations for both RTs and error rates. Note that I only report statistics for the congruency effect as well as for effects involving the relevant factor experiment within the between-experiment comparison. In addition to the between-experiment analysis, Experiment 7b was also analyzed in the same way as Experiment 7a.

### 4.3.2 RESULTS

The exclusion of all non-contingent trials of the reduced contingency group (equivalent to 90 trials out of 450 trials per participant) resulted in 19,440 trials to be analyzed. After applying the cleansing procedure as described for Experiment 2–6 (see *Section 3.2.2*), I excluded 461 trials corresponding to 2.4% of the 19,440 trials.

### 4.3.2.1 Between-experiment comparison

When looking at saccade latencies, the between-experiment comparison revealed a robust effect of congruency across experiments,  $F(2, 176) = 19.46$ ,  $\varepsilon = .92$ ,  $p < .001$ ,  $\eta^2_p = .18$ . The main effect of experiment was not significant,  $F < 1$ , and there were no significant interactions involving this factor,  $ps > .05$ .

The between-experiment analysis of error rates also showed a significant congruency effect,  $F(2, 176) = 8.90$ ,  $\varepsilon = .89$ ,  $p < .001$ ,  $\eta^2_p = .09$ , which is further qualified by the interaction of congruency and experiment,  $F(2, 176) = 4.62$ ,  $\varepsilon = .89$ ,  $p = .014$ ,  $\eta^2_p = .05$ . Pairwise comparisons indicated that the congruency effect (in terms of significantly lower error rates in the congruent vs. incongruent condition) was present in Experiment 7a,  $p = .001$  for the comparison between congruent and incongruent condition, while this difference was not significant in Experiment 7b,  $p = .077$ . However, the result pattern of the unrelated condition was also different for Experiment 7a versus Experiment 7b. In the former, the error rate of the unrelated condition was in between that of the congruent and incongruent conditions (and closer to the incongruent condition), while the error rate of the unrelated condition in Experiment 7b was lower than the error rates in both the congruent and incongruent conditions (and closer to the congruent condition). The main effect of experiment was not significant,  $F < 1$ , and the same held for all remaining interactions involving this factor,  $ps > .05$ .

### 4.3.2.2 Temporal oculomotor performance in Experiment 7b

The ANOVA yielded a significant main effect of congruency,  $F(2, 88) = 10.15$ ,  $p < .001$ ,  $\eta^2_p = .19$ . As in Experiment 7a, participants initiated their saccades faster in the congruent condition ( $M = 251$  ms,  $SE = 6$  ms) compared to both the incongruent condition ( $M = 259$  ms,  $SE = 7$  ms;  $p = .003$  for the pairwise comparison) and the unrelated condition ( $M = 259$  ms,  $SE = 7$  ms;  $p < .001$ ). RTs did not differ significantly between the incongruent and the unrelated condition,  $p = .595$ . The interaction of contingency and congruency failed

to reach significance,  $F(2, 88) = 2.44, p = .093, \eta^2_p = .05$ , even though the observed data pattern pointed in the predicted direction. The effects of contingency, of context, and the interaction of congruency and context did not approach significance, all  $F_s < 1$ . Further, neither the interaction of contingency and context,  $F(1, 44) = 1.17, p = .284, \eta^2_p = .03$ , nor the three-way interaction,  $F(2, 88) = 1.77, p = .176, \eta^2_p = .04$ , were significant.

#### 4.3.2.3 Oculomotor error rates in Experiment 7b

Participants showed, on average, error rates of 7.9 % ( $SE = 0.8\%$ ) in the congruent condition, 8.9 % ( $SE = 0.9\%$ ) in the incongruent condition, and 7.2% ( $SE = 0.7\%$ ) in the unrelated condition, yielding a significant main effect of congruency,  $F(2, 88) = 4.88, \varepsilon = .88, p = .013, \eta^2_p = .10$ . Pairwise comparisons revealed a significant difference between the incongruent and unrelated condition,  $p = .006$ . Though there was a numerical tendency that error rates were lower for the congruent versus incongruent condition, pairwise comparisons showed that this difference was not statistically significant,  $p = .095$ . Furthermore, error rates did not differ between the congruent and unrelated condition,  $p = .132$ . All other statistical comparisons were not significant, all  $F_s < 1$  (except for the interaction of congruency and context,  $F(2, 88) = 1.96, \varepsilon = .88, p = .152, \eta^2_p = .04$ , and the three-way interaction,  $F(2, 88) = 1.20, \varepsilon = .88, p = .303, \eta^2_p = .03$ ).

#### 4.3.2.4 Exploratory analyses in Experiment 7b

##### *Speed analysis*

Again, I observed a (self-evident) effect of speed on saccade latencies in Experiment 7b,  $F(1, 44) = 365.63, p < .001, \eta^2_p = .89$ . The interaction of congruency and speed was not significant,  $F(1, 44) = 1.44, p = .237, \eta^2_p = .03$ , as well as all other effects involving the factor speed, all  $F_s < 1$ .

When analyzing error rates with respect to speed in Experiment 7b, a significant main effect of speed indicated that participants' reactions were more error-prone at fast RTs

( $M = 11.8\%$ ,  $SE = 1.2\%$ ) compared to slow RTs ( $M = 5.1\%$ ,  $SE = 0.5\%$ ),  $F(1, 44) = 50.08$ ,  $p < .001$ ,  $\eta^2_p = .53$ . The interaction of congruency and speed was not significant,  $F(1, 44) = 3.39$ ,  $p = .072$ ,  $\eta^2_p = .07$ . The three-way interaction of congruency, context, and speed,  $F(1, 44) = 1.90$ ,  $p = .175$ ,  $\eta^2_p = .04$ , and the four-way interaction,  $F(1, 44) = 1.44$ ,  $p = .236$ ,  $\eta^2_p = .03$ , were not significant, and the same held for all other effects involving the factor speed, all  $F_s < 1$ .

#### *Initial performance analysis*

The analysis of initial saccade latencies revealed no significant congruency effect,  $F < 1$ , and also the interaction of congruency and context failed to reach significance,  $F(1, 44) = 3.70$ ,  $p = .061$ ,  $\eta^2_p = .08$ . The main effect of context,  $F(1, 44) = 1.94$ ,  $p = .171$ ,  $\eta^2_p = .04$ , and the interaction of context and contingency,  $F(1, 44) = 1.06$ ,  $p = .309$ ,  $\eta^2_p = .02$ , were not significant. None of the remaining effects approached significance, all  $F_s < 1$ .

When initial error rates were analyzed, the congruency effect was not significant,  $F < 1$ . Also, the effect of context on error rates was not significant,  $F(1, 44) = 3.38$ ,  $p = .073$ ,  $\eta^2_p = .07$ . There was neither a significant effect of contingency,  $F(1, 44) = 1.68$ ,  $p = .202$ ,  $\eta^2_p = .04$ , nor a significant three-way interaction,  $F(1, 44) = 1.31$ ,  $p = .258$ ,  $\eta^2_p = .03$ . None of the remaining effects reached significance, all  $F_s < 1$ .

#### *Temporal dynamics analysis*

Due to a single software error causing missing readouts from the eye tracker for Block 5 for one participant, the corresponding data set of this participant had to be excluded for the block analysis of saccade latencies and error rates. Saccade latencies were submitted to a four-way ANOVA with block as an additional factor. A significant influence of block,  $F(4, 172) = 6.32$ ,  $\varepsilon = .78$ ,  $p < .001$ ,  $\eta^2_p = .13$  could be observed. Participants responded after 264 and 263 ms ( $SE = 8$  ms and 8 ms) in Block 1 and Block 2, respectively, whereas they were significantly faster in Block 3–5 ( $M = 249/247/249$  ms,  $SE = 7 / 7 / 9$  ms). Pairwise

comparisons revealed that this effect of block mainly resulted from the decrease in saccade latencies between Blocks 2 and 3,  $p = .001$ , whereas saccade latencies did not differ between Blocks 1 and 2,  $p = .804$ , and between Blocks 3–5,  $ps > .050$ . Further, block and context interacted significantly,  $F(4, 172) = 3.18$ ,  $\varepsilon = .78$ ,  $p = .025$ ,  $\eta^2_p = .07$ . Separate repeated-measures ANOVAs for each context condition showed that saccade latencies decreased over time for the social context condition,  $F(4, 88) = 5.31$ ,  $\varepsilon = .74$ ,  $p = .003$ ,  $\eta^2_p = .19$ , but did not vary as a function of block for the non-social context condition,  $F < 1$ . All remaining interactions involving the factor block were not significant, neither the interaction of block and congruency,  $F(4, 172) = 1.25$ ,  $\varepsilon = .82$ ,  $p = .296$ ,  $\eta^2_p = .03$ , the interaction of block and contingency,  $F(4, 172) = 1.77$ ,  $\varepsilon = .78$ ,  $p = .153$ ,  $\eta^2_p = .04$ , nor any of the other interactions, all  $F_s < 1$ .

The block-wise error rate analysis showed no significant effect involving the factor block (all  $F_s < 1$ , except for the main effect of block,  $F(4, 172) = 1.47$ ,  $p = .215$ ,  $\eta^2_p = .03$  and the interaction of block and context,  $F(4, 172) = 1.15$ ,  $p = .333$ ,  $\eta^2_p = .03$ ).

### 4.3.3 DISCUSSION

By comparing the results of Experiments 7a and 7b, I intended to test whether explicit versus implicit knowledge about the action-effect contingencies alters action-effect learning in conjunction with perfect versus reduced action-effect contingency. This between-experiment analysis revealed two main findings. First, the between-experiment analysis did not reveal any effects of experiment in RTs, indicating the absence of any meaningful differences in the observed result patterns between Experiments 7a and 7b. In principle, this observation is in line with the results of Elsner and Hommel (2004), who showed that effects of action-effect learning are comparable in size, regardless of whether participants were asked to attend to the action-effect relationship or not. However, note that the experiments com-

pared with each other by Elsner and Hommel differed not only with respect to the instruction, but also in that *contiguity* was manipulated in one experiment, while *contingency* was manipulated in the experiment where participants received explicit instructions (Elsner & Hommel, 2004). Further, participants in their study received the instruction to attend to the action-effect relationship rather than obtaining explicit knowledge about the nature of this relationship.

Second, the analysis of error rates showed that the explicit knowledge about the action-effect contingencies affected error rates such that participants committed fewer errors in the congruent as compared to the incongruent condition. This difference in error rates was especially pronounced in Experiment 7a, and there was a numerical trend evident into the same direction in error rates of Experiment 7b, while this difference in error rates was not significant in Experiment 7b. It might be that the provided knowledge about the action-effect contingencies induced some kind of increased awareness for the relation between the to-be-executed action and the resulting action effect, which in turn reduced effects of distraction as caused by the presentation of the task-irrelevant stimulus. However, this effect was only evident in error rates, not in RTs, presumably because RTs in the presented results represent the more sensitive dependent measure, which is less easily affected by any kind of increased awareness.

The analyses of both Experiments 7a and 7b revealed that the congruency effect in saccade latencies (and error rates) was unaffected by the contingency manipulation. Several explanations appear conceivable: First, the results might reflect specific mechanisms inherent in oculomotor action control. That is, oculomotor action-effect learning might be rather robust against reduced predictability of the upcoming effect – in stark contrast to manual action-effect learning (cf. Elsner & Hommel, 2004). This deliberation seems reasonable considering the fact that the “simplest” oculomotor action effect, that is, the perception of the post-saccadic object to perceive relevant new information, is quite often subject to reduced

---

predictability because of acuity limitations in peripheral vision (Land & Tatler, 2009, but see Herwig & Schneider, 2014, for an explanatory framework of how humans are able to overcome these acuity limitations). On a more applied level, one could imagine that the foundations of our environment, such as buildings or landscapes, remain constant, while there are many sources of possible environmental change. For example, if you look out of a window, you may see a cloudless blue sky. A few minutes later, however, the same view from the window can reveal a cloudy sky. Thus, the repeated experience that saccadic input is not perfectly predictable might contribute to a lack of influence of the contingency manipulation.

Second, the absence of any influence of contingency might indicate that the contingency manipulation as implemented within both Experiments 7a and 7b was too weak to be reflected in the dependent measures. Likely, the effects of reduced contingency would be evident if I would have further reduced the ratio of contingent to non-contingent trials, for example to a ratio of 70% of contingent to 30% of non-contingent trials (cf. Dignath et al., 2018, where reduced contingency significantly affected social affiliation judgements during imitation at a ratio of 75% of contingent to 25% of non-contingent trials). If the contingency drops close to 50%, however, no action-effect learning would be possible any longer in the first plan, as the effects would no longer be predictable at all. If one compares the compatibility manipulation implemented in Experiments 7a and 7b with the manipulation implemented in the study by Elsner and Hommel (2004), differences in the degree of granularity of the contingency manipulation are noticeable. In this previous study, the authors manipulated the contingency between key presses and subsequent tones in five different ways by varying the relative frequencies of the presence or absence of an effect tone given the presence or absence of the corresponding key press. Evidence in favor of action-effect learning was observed only when the effect tone occurred more often in the presence than in the absence of the action and when the overall frequency of the effect sound was high (Elsner

---

& Hommel, 2004). No indication of an acquisition of action-effect associations was, for example, evident in a group where the overall frequency of the effect tone was low (i.e., when an effect tone was played in only 10% of the test phase trials). Thus, the contingency manipulation in Experiments 7a and 7b differed from the manipulation as implemented by Elsner and Hommel with respect to the fact that a saccade was requested in every trial here (vs. no-go trials without any action demands which might or might not trigger an effect tone in that previous study), and that every saccade toward any of the two targets resulted in a visual action effect here (vs. conditions without any action effects in that previous study). Thus, it might be possible that a more granular contingency manipulation would demonstrate that the contingency between action and effect could indeed influence action-effect learning also in the oculomotor domain.

Even though the between-experiment comparison suggested a rather comparable result pattern across Experiments 7a and 7b, it is worth taking a closer look at the isolated results of Experiment 7b. In line with ideomotor theory, the analysis of RTs demonstrated a congruency effect, just as observed in Experiments 3, 5, and 7a. Although not statistically significant, there was a numerical trend in the data that the congruency effect was smaller for the reduced as compared to the perfect contingency group (see *Figure 12*). This finding is interesting, since it reflects the results pattern as hypothesized, but not observed, for Experiment 7a, where no instructions about action-effect contingencies were provided. In Experiment 7b, however, I hypothesized congruency effects to be less affected (or even not at all) by the contingency manipulation since participants were explicitly told about these contingencies. That is, rather than observing benefits of explicit knowledge about the action-effect contingencies, the present results show – if anything – rather a trend toward less intensive action-effect learning in the reduced contingency group when explicitly instructed about the action-effect contingencies. However, this observation only reflects a numerical trend, which



---

must be interpreted with caution and requires further empirical clarification, for example by broadening the range of the contingency manipulation.

Like in Experiment 7a, I conducted some exploratory analyses to specify the nature of the congruency effect. Most importantly, these analyses provided consistent results with respect to the contingency manipulation: In brief, none of these analyses showed any influence of the contingency manipulation on the observed results. The speed analysis showed the same speed-accuracy trade-off as in Experiment 7a: Participants were more error-prone in trials where they responded rather fast as compared to trials where they took more time to initiate the target saccade. However, and in contrast to Experiment 7a, there was no interaction of response speed and congruency in error rates, which fits with the fact that the influence of congruency on error rates was generally rather weak in Experiment 7b. When analyzing the first few trials only, there were neither congruency effects evident in RTs nor in error rates, independent of the contingency or context manipulation, similar to Experiment 7a (see *Section 4.2.3* for a more elaborate discussion of this finding). The analysis of the temporal dynamics of the congruency effect demonstrated decreasing RTs over the course of the experiment. However, note that the temporal dynamics analysis of Experiment 7b was the only analysis where the factor context had a significant influence on saccade latencies. The interaction of block and context indicated that RTs only decreased over the course of the experiment for the social context condition, while RTs remained at the same level throughout the experiment for the non-social context condition. I can only speculate about possible reasons for this observation, but it might be that the faces as presented in the social context initially required deeper and more intense visual processing as compared to the black-and-white ellipses used as stimuli in the non-social condition. Therefore, RTs were higher for a social (vs. non-social) context at the beginning of the experiment, while the corresponding increased processing effort subsided over the course of the experiment due to the repeated presentation of the faces.

Taken together, the results of Experiments 7a and 7b provide strong evidence for the acquisition of associations between performing a saccade and perceiving the subsequent visual action effect as reflected in robust congruency effects in saccade latencies. Contrary to what I expected, this congruency effect in saccade latencies was unaffected by the contingency manipulation, that is, participants picked up these action-effect regularities in a similar manner independent of whether action-effect contingency was perfect or reduced. As a contingency reduction of 20% can be considered quite substantial, this result can be interpreted in terms of an astonishing resistance of the formation of associations between action and effect against imperfections and disturbances that one would probably not expect based on the associative learning literature. Finally, these main results have shown to be independent of a social versus non-social context.

---

## 4.4 EXPERIMENT 8

So far, the literature on temporal aspects of action-effect learning focused predominantly on the investigation of the temporal contiguity between action and effect using arbitrary action-effect associations in *non-social* environments (e.g., Dignath et al., 2014, for the manual, and Experiment 1 within this work for the oculomotor domain). In these studies, participants experienced that (manual or oculomotor) actions produced specific effects that occurred with a short or long temporal delay after their actions, that is, with high or low temporal contiguity.<sup>14</sup> High contiguity was shown to be beneficial for the acquisition of associations between motor actions and their subsequent perceptual effects (Elsner & Hommel, 2004; but see Dignath et al., 2014, for divergent results). In general, social action effects often tend to come with reduced temporal contiguity compared to most action effects in the non-social, inanimate environment (see *Section 2.1.3* for a detailed discussion on the peculiarities of social action effects). Up to now, it is not yet clear whether this naturally occurring delay between an agent's action and the response brought about by another agent is reflected in more or less effective action-effect learning with shorter versus longer action-effect intervals. A direct comparison of how short versus long action-effect intervals affect action-effect learning differentially across a social versus non-social context is still pending.

Experiment 8 aimed at filling this gap by investigating the role of contiguity on action-effect learning in a direct comparison of social versus non-social action effects. To this end, the experimental design of Experiments 2–6 was slightly adapted. Again, participants acquired novel oculomotor action-effect associations by alternately looking at two schematic faces on the screen that consistently responded to the participant's gaze with either

---

<sup>14</sup> Note that the terminology action-effect interval (AEI) and action-effect delay are used as synonyms throughout this work.

direct or averted gaze. I relied on the same congruency manipulation as implemented in Experiments 2–6 to test whether anticipation of the partner’s gaze response occurs prior to saccade preparation. For the non-social equivalent of the faces, the same white ellipses as in Experiment 5 were presented. Looking at these ellipses consistently triggered the appearance of a black square either centrally or laterally within the ellipsis. The present experimental design differed from the design of Experiments 2–6 with respect to the temporal onset of the visual action effect. While the action effect appeared on the screen with a constant AEI of 200 ms in Experiments 2–6, the temporal onset of the action effect was either immediate (0 ms AEI) or delayed (200/600 ms AEI) in the present experiment.

Experiment 8 was designed to address two research questions. First, I asked which temporal delay is most effective in developing oculomotor action-effect associations across social and non-social contexts. Based on the fact that social action effects are often less immediate as compared to non-social effects, I predicted enhanced action-effect learning with more delayed effects in a social context, whereas immediate effects may be more suitable to acquire action-effect associations with non-social effect signals. More specifically, I expected congruency effects to be larger for long (vs. short) delays in a social context, while this should be the other way round in a non-social context. Second, I tested whether the previously reported dependency of saccade latencies on the AEI between action and effect (Experiment 1) can be replicated and extended for a broader spectrum of temporal intervals. Since it was demonstrated that saccade latencies are significantly affected by temporal AEIs that only differ by 200 ms in Experiment 1 (see *Section 3.1*), I expected to observe substantial differences in saccade latencies when the spacing between the different delay conditions is even greater.

---

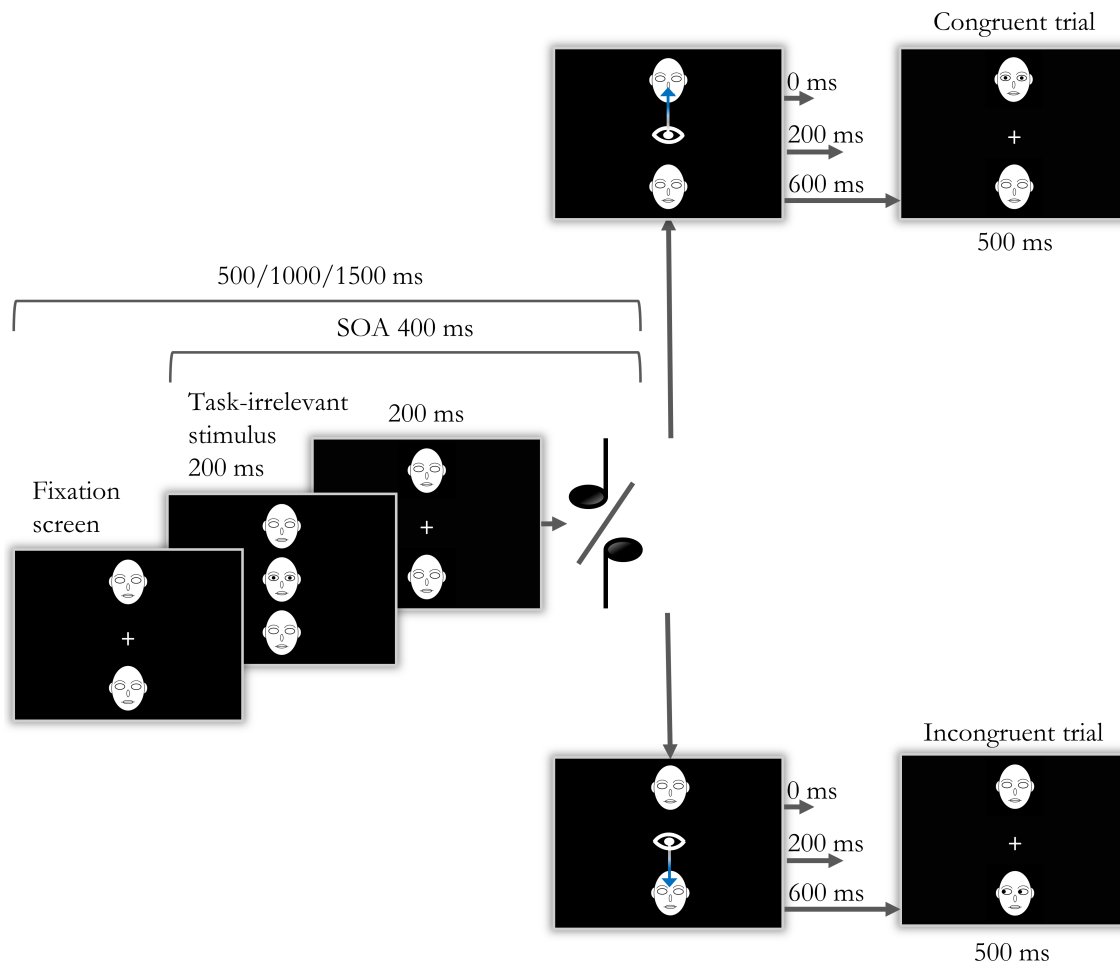
## 4.4.1 METHODS

### 4.4.1.1 Participants

Seventy-six participants were recruited for this experiment, who received either course credits or monetary compensation for participation. Participants were naïve with respect to the purpose of the experiment and gave written informed consent before the study started. All participants reported normal or corrected-to-normal vision. The data of three participants were excluded from analysis due to an unusually high error rate ( $> 30\%$  in at least one cell of the design). One additional participant was excluded from analysis, since there were not sufficiently valid data in each cell of the design after data cleaning. Data of the remaining 72 participants (mean age = 24.1 years,  $SD = 3.9$  years, age range: 18 – 36 years, 21 male) were analyzed.

### 4.4.1.2 Stimuli and procedure

The experimental procedure was the same as in Experiments 2–6 with the crucial difference that I manipulated the AEI between action and effect (see *Figure 13*), while this timing was held constant in Experiments 2–6. More specifically, the AEI at which the targeted baseline face changed into direct versus averted gaze was manipulated between subjects: The change either occurred immediately with the saccade entering the target area (i.e.,  $AEI = 0$  ms), or with a delay of 200 ms or 600 ms after the saccade entered the target area. As in Experiments 7 and 8, the SOA manipulation has been discarded in the present experiment. The experiment lasted around 35 minutes, and consisted of 15 experimental blocks with 30 trials each (450 trials in total). A calibration of the eye tracker was performed prior to each experimental block. Trial order was pseudo-randomized (maximum sequence length of the same imperative tone stimulus was restricted to seven).



*Figure 13.* Illustration of the trial structure in Experiment 8 for schematic face stimuli. Each trial started with the presentation of a black screen with a white fixation cross and two peripheral baseline face stimuli (without gaze). After that, the fixation cross was replaced by a task-irrelevant stimulus, which was either a baseline face stimulus (without any visible gaze) or a face with direct or averted gaze. Then, an auditory stimulus (high vs. low) signaled the participant to respond with a corresponding gaze shift (up vs. down). After the saccade had entered the upper or lower face target area, either a direct or an averted gaze was displayed (contingent upon the targeted face) after a temporal delay (action-effect interval, AEI) of 0/200/600 ms for a duration of 500 ms, before the next trial started. The central visual stimulus was either congruent, incongruent, or unrelated to the subsequent gaze response of the targeted face. The trial sequence shown in this figure depicts a congruent and incongruent trial sequence with a direct gaze as the task-irrelevant stimulus. In case of an averted gaze as task-irrelevant stimulus (not depicted in the present figure), a gaze shift toward the upper target face (triggering direct gaze) would represent an incongruent trial, and a gaze shift toward the lower target face (triggering averted gaze) would represent a congruent trial. For the unrelated congruency condition, the task-irrelevant stimulus was a baseline face without any visible gaze.

#### 4.4.1.3 Design and analysis

Data were submitted to a repeated-measures ANOVA with *congruency* (congruent vs. incongruent vs. unrelated) as within-subjects factor, and *AEI* (0 ms vs. 200 ms vs. 600 ms) as well as *context* (social vs. non-social) as between-subjects factors.

## 4.4.2 RESULTS

Applying the same data cleansing procedure as in Experiments 2–6 resulted in the exclusion of 959 trials in total, which corresponded to 3.0% of 32,400 trials in total.

### 4.4.2.1 Temporal oculomotor performance

Saccade latency amounted to 251 ms ( $SE = 7$  ms) for congruent trials, to 254 ms ( $SE = 7$  ms) for incongruent trials, and to 256 ms ( $SE = 7$  ms) for unrelated trials, and these saccade latencies differed significantly,  $F(2, 132) = 3.48, p = .034, \eta^2_p = .05$  (see *Figure 14*). RTs were significantly faster in congruent trials compared to unrelated trials,  $p = .014$ , and there was a numerical tendency of participants responding faster in congruent as compared to incongruent trials,  $p = .092$ . RTs did not differ significantly between incongruent and unrelated trials,  $p = .407$ . The effect of congruency was further specified by the significant interaction of congruency and AEI,  $F(4, 132) = 2.68, p = .035, \eta^2_p = .08$ . For further analysis of the significant interaction, separate one-way ANOVAs for each AEI condition were computed. These analyses revealed that congruency affected saccade latencies in the AEI = 200 ms condition,  $F(2, 46) = 6.10, p = .004, \eta^2_p = .21$ , but neither in the 0 ms condition,  $F(2, 46) = 2.18, p = .125, \eta^2_p = .09$ , nor in the 600 ms condition,  $F < 1$ . In the 200 ms condition, saccade latencies were (almost) significantly faster for congruent compared to incongruent trials,  $p = .052$ , and to unrelated trials,  $p = .004$ , while incongruent and unrelated trials did not significantly differ from each other,  $p = .130$ . No other effect approached significance, neither the effect of AEI,  $F < 1$ , nor the effect of context,  $F(1, 66) = 1.28, p = .262, \eta^2_p = .02$ , nor the interaction of AEI and context,  $F(2, 66) = 1.35, p = .267, \eta^2_p = .04$ , nor any of the remaining interactions, all  $F$ s  $< 1$ .

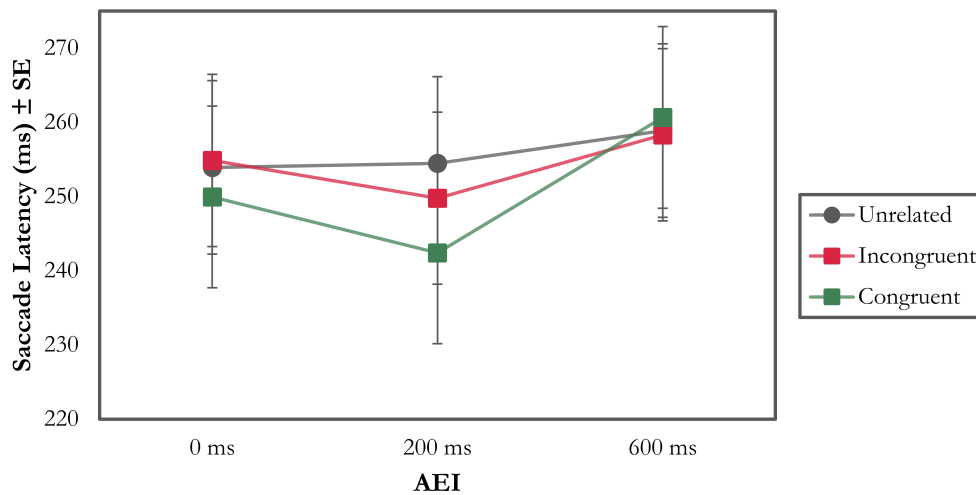


Figure 14. Mean saccade latencies (in ms) in Experiment 8 as a function of congruency (congruent vs. incongruent vs. unrelated) and action-effect interval (AEI: 0 ms vs. 200 ms vs. 600 ms). Error bars represent standard errors of the mean ( $SE$  in ms).

#### 4.4.2.2 Oculomotor error rates

The ANOVA revealed a significant effect of congruency,  $F(2, 132) = 5.98$ ,  $\varepsilon = .89$ ,  $p = .005$ ,  $\eta^2_p = .08$ . Paired  $t$ -tests indicated a tendency towards fewer errors in the congruent (3.8%,  $SE = 0.3\%$ ) compared to the incongruent condition (4.3%,  $SE = 0.4\%$ ),  $p = .066$ . The error rate was smallest for the unrelated condition (3.3%,  $SE = 0.3\%$ ), and significantly differed between the unrelated and the incongruent conditions,  $p = .003$ , while the congruent and unrelated conditions did not significantly differ,  $p = .067$ . There was no significant effect of context,  $F(1, 66) = 2.27$ ,  $p = .137$ ,  $\eta^2_p = .03$ , no significant effect of AEI,  $F < 1$ , and no significant interactions, all  $F$ s  $< 1$  (except for the interaction of context and AEI,  $F(2, 66) = 2.64$ ,  $p = .079$ ,  $\eta^2_p = .08$ ).

#### 4.4.3 DISCUSSION

Experiment 8 assessed which temporal action-effect interval (AEI) is most effective for the build-up of oculomotor action-effect associations in a social versus non-social envi-



---

ronmental context. Therefore, I relied on an adapted version of the paradigm used in Experiments 2–6, where congruency effects served as an empirical marker of action-effect learning, and additionally manipulated the temporal effect delay.

First, the results demonstrated congruency effects in terms of faster saccade latencies and reduced error rates for the congruent as compared to the incongruent condition, with the restriction that the congruency effect in saccade latencies was only evident in the AEI = 200 ms condition. This result represents a further replication of the central ideomotor effect and thereby demonstrates its robustness. Moreover, the three-way interaction of congruency, contiguity and context was not significant, neither in RTs nor in error rates, rejecting the hypothesis that the strength of the acquired action-effect associations should depend on the AEI in conjunction with a social or non-social context. This hypothesis about a different influence of contiguity as a function of a social versus non-social context was based on theoretical deliberations (see Kunde et al., 2018), claiming that social action effects likely occur with greater temporal delay than non-social effects. For example, the auditory effect of knocking on a door occurs immediately after the action, while it takes some time for another person to respond accordingly when being waved at. However, counter-examples show that delayed action effects are not exclusively reserved for animate environments. Imagine that you click on a hyperlink in the web browser when the internet connection is poor: It will take some time until the requested website is loaded. Alternatively, and especially for the oculomotor domain, when you unlock your smartphone using face recognition, it will take a few seconds for the phone to process your face before you can finally use it. These valid examples showing the existence of delayed intended effects in inanimate environments demonstrate that the lack of an interacting influence of the context and the contiguity manipulation in the present Experiment 8, while contradicting the initial hypothesis, still appears reasonable in the light of these considerations.

However, the assumption that both social and non-social action effects can principally kick in immediately or delayed in daily life cannot account for the fact that congruency effects in saccade latencies did not significantly occur across all AEI conditions, independent of the type of environment (social vs. non-social). A congruency effect in saccade latencies significantly occurred in the AEI = 200 ms condition, and there was a numerical trend evident (albeit not significant) towards such an effect in the AEI = 0 ms condition, while there was no such effect in the AEI = 600 ms condition. The congruency effect in error rates, however, was present across all AEI conditions. These findings support the observation that congruency effects, as reported in the previous Experiments 2–7 of this work, were always present with a constant AEI of 200 ms. Further, these findings are in line with the results of Experiment 1, where the acquisition of action-effect associations has been shown for all three AEI conditions, albeit weakened for the 0 ms as compared to the two delayed (100/300 ms) conditions (see *Section 3.1*).

The weakened evidence for action-effect learning in the immediate AEI condition in Experiment 1 caused me to adjust the corresponding condition in the present experiment. In order to increase the saliency of the visual action effect in the AEI = 0 ms condition and to minimize potential saccadic suppression phenomena, the design of the AEI = 0 ms condition of Experiment 8 differed from that of the respective condition in Experiment 1. While the onset of the visual change was determined by the first saccade leaving the interest area around the fixation cross in Experiment 1, which implies that the change happened *during* the target saccade, the action effect occurred with the saccade *entering* the target area in Experiment 8. For the AEI = 0 ms condition, a significant congruency effect was observed in error rates, but the corresponding effect only occurred as a numerical trend in saccade latencies. This result suggests that the saliency of the action effect in the AEI = 0 ms condition of the present experiment was still not strong enough to form effective action-effect associations so that the presentation of the action effect as a task-irrelevant stimulus could have

---

triggered the associated action thereby affecting RTs in a significant manner. It might be that action-effect learning is weakened with immediate delays, and only spatial saccade parameters as measured in Experiment 1, or error rates as measured in this experiment, but not saccade latencies, are capable of revealing these subtle learning processes.

With respect to the AEI = 600 ms condition, the absence of any congruency effects might be due to a reduced feeling of agency, which is the subjective feeling of being the cause of a specific action and the action-associated effects (Gallagher, 2000; Haggard, 2005). Agency judgements are moderated by variables containing information about a potential causal relationship, like temporal contiguity between action and effect. For example, agency ratings were reduced for delayed relative to immediate action effect tones (Weller, Schwarz, Kunde, & Pfister, 2017). If participants did not feel (or only felt reduced) agency over the visual action effect in the AEI = 600 ms condition of the present experiment, they might not have paid much attention toward these action effects. Therefore, participants might have failed to build up sufficiently strong action-effect associations, and the congruency manipulation was therefore ineffective. However, this post-hoc explanation is speculative and requires further clarification, for example by additionally collecting agency ratings within the present paradigm.

A second set of results pertains to the previously reported dependency of saccade latencies on the AEI (see Dignath et al., 2014, for the manual domain, and Experiment 1 for the oculomotor domain). In the present experiment, however, participants did not show any systematic differences in saccade latencies as a function of the AEI. This finding contradicts the findings of Experiment 1 (see *Section 3.1*), where an even smaller variation of the AEI (0/100/300 ms in Experiment 1 vs. 0/200/600 ms in the present experiment) was reflected in saccade latencies, and of further corroborating evidence on the integration of action-specific effect delays into action control from the manual domain (e.g., Dignath et al., 2014;

Kiesel & Hoffmann, 2004; Kunde, 2003). Thus, this raises the question of whether participants in the present experiment did not integrate the information about the temporal delay between action and effect into an action-delay-effect association (as suggested by Dignath et al., 2014), or whether they built such associations, but these failed to affect saccadic initiation times due to the nature of the experimental procedure of the present experiment. In fact, the experimental procedure here differs from that in Experiment 1 with respect to several aspects, among which procedural differences prior to saccade execution might be the most relevant ones. While participants were free to decide which target to saccade to after onset of the targets on the screen in Experiment 1, the procedure of the present experiment involved the presentation of a task-irrelevant prime stimulus and an imperative auditory stimulus which defined the saccade target. It might be that this additional cognitive load prior to saccade execution in the present experiment prevented the action-delay-effect associations from coming into effect. However, this hypothesis is of course speculative and needs empirical validation – primarily by replicating and extending the results from Experiment 1 to different AEIs and stimulus material – before any further conclusions can be drawn.

A study of Pfister et al. (2013) addressed the impact of anticipated social action effects in an imitation setting where being imitated, rather than counter-imitated, has been shown to facilitate motor actions. In their study, participants worked in pairs – one representing the *model* and the other one representing the *imitator*. The model was instructed to respond to a color cue with either a short or a long key press. In different blocks, the imitator responded either with imitation (i.e., with a congruent action effect) or with counter-imitation (i.e., with an incongruent action effect). That is, the imitator performed a short (long) key press in response to a short (long) key press of the model in the congruent (incongruent) condition. The authors observed that the model's response was faster when being imitated (congruent condition) rather than being counter-imitated (incongruent condition), suggesting that the anticipated response of the imitator served as mental cue for the model to initiate her/his

---

own key press (for similar findings, see also Pfister et al., 2017; Weller, Pfister, & Kunde, 2019). Interestingly, and relevant for the current discussion, the imitator also responded faster while imitating rather than counter-imitating the model, suggesting that the anticipation of fast versus slow responses of the imitator, that is, of short versus long action-effect delays, might at least partially contribute to explaining the observed effects on the model's side. Thus, Pfister and colleagues (2017) additionally manipulated the temporal delay between the model's action and the imitation response to disentangle whether previously reported imitation effects are driven by action-effect compatibility or action-effect delay. In contrast to action-effect compatibility, the action-effect delay (short vs. long) had no influence on action initiation in their study (Pfister et al., 2017). In a similar vein, results from Lelonkiewicz and Gambi (2017) and Lelonkiewicz, Gambi, Weller, and Pfister (2020) also could not find convincing evidence for an integration of temporal information into action planning in a social imitation setting.

Together with the findings of the present experiment, the absence of a notable effect of the temporal delay on action planning as reported above might be explained in the light of the sociomotor framework of action control (Kunde et al., 2018), which suggests that effect delays could be less likely integrated into action planning for social versus non-social action contexts, since social interaction is characterized by more variable effect delays. However, several findings speak against a different influence of action-effect delays depending on a social versus non-social context. First, there is convincing evidence that anticipation of other people's timing indeed plays a role in social interactions (Knoblich & Jordan, 2003; Pecenka & Keller, 2011; Wolf, Sebanz, & Knoblich, 2018). Second, the results of the non-social context condition of Experiment 8 of the present work provide an example where effect delays did not seem to be incorporated into action planning even for a non-social context. Together, these examples contradict the possibility that the absence of an influence of the AEI on action initiation is a specific feature of social action control, and rather point

toward the necessity for further empirical work to investigate the role of action-effect contingency in both social and non-social context conditions using different tasks and a broad range of action-effect delays.

## 4.5 CONCLUSION OF CHAPTER 4

Taken together, the results of Experiments 7a, 7b, and 8 again provide evidence for the acquisition of associations between performing a saccade and perceiving the subsequent visual action effect as reflected in strong (Experiments 7a and 7b) and slightly weaker (Experiment 8) congruency effects in saccade latencies, using a similar setup as first implemented by Huestegge and Kreutzfeldt (2012). Thus, based on the experiments conducted so far, the effect can be considered very stable.

Contrary to what I expected, this congruency effect in saccade latencies was unaffected by contingency, that is, participants did not pick up these action-effect regularities in a fundamentally different manner when action-effect contingency was perfect versus reduced.

Congruency effects were evident in error rates independent of contiguity. With respect to RTs, the contiguity manipulation affected action-effect learning such that the congruency effect was most pronounced in the 200 ms delay condition, while there was a numerical trend toward a congruency effect in the 0 ms delay condition, but no congruency effect in the 600 ms delay condition. That is, the results suggest that congruency effects disappear with greater delays, but further investigation is necessary to specify the range, which is effective for the build-up of effective action-effect associations. Further, the present data provide no indication for the assumption that the temporal delay generally becomes an integral part of the action representation that affects action generation. Potential reasons for the observed results (e.g., with respect to the specific role of the oculomotor system or the

---

quality of the manipulations) and implications for further investigations have already been laid out in the respective discussion parts of Experiments 7a, 7b, and 8.

Most notably, no evidence for a specific role of a social (vs. non-social) context was found in the main analyses of Experiments 7a, 7b, and 8, thereby in general replicating the findings of Experiments 2–6. This absence of meaningful differences for a social versus non-social context suggests that sociomotor learning in the oculomotor domain follows the same principles as in general (non-social) action control contexts (cf. Experiments 3 and 5). However, it might also mirror a failure to establish a more valid “social” context, which allows capturing “real” social behavior comparable to behavior in situations with real human interaction partners. Since this question is relevant for the entire work, I will further elaborate on this issue in the General discussion (see *Section 7.3.1*).

## 5 THE ROLE OF THE ACTION MODE DURING ACQUISITION

### 5.1 ACQUISITION AND EFFECTS OF ACTION-EFFECT ASSOCIATIONS IN THE MANUAL DOMAIN

Numerous studies on the acquisition of action-effect associations have referred to the two-phase action-effect acquisition paradigm (see *Section 2.1.2.1*), implementing either free-choice or forced-choice test phases, or both (Dignath et al., 2014; Elsner & Hommel, 2001; Herwig et al., 2007; Herwig & Waszak, 2009; Hoffmann et al., 2009; Hommel et al., 2003). Across them, and irrespective of the particular focus of each study, reliable congruency effects were obtained for both free-choice and forced-choice test phase designs.

The original study of Elsner and Hommel (2001) was designed such that participants could freely select one of two response alternatives in the acquisition phase. The majority of subsequent studies investigating action-effect associations transferred the free-choice acquisition phase into their respective designs. The question of whether the action mode during the acquisition phase has an impact on congruency effects in the test phase was first addressed by Herwig et al. (2007) and Herwig and Waszak (2009). They additionally incorporated a forced-choice acquisition phase into the action-effect acquisition paradigm. An imperative stimulus (either arbitrary stimuli, e.g., colored shapes, or stimuli spatially compatible to the required responses, e.g., arrows pointing toward the left or right) indicated which of the two responses (left or right key press) should be performed. Analogous to the



---

free-choice acquisition, each key press contingently produced a certain effect tone. The subsequent forced-choice test phase was the same for both types of acquisition groups. The typical congruency effect was exclusively observed for the subgroup that experienced a free-choice acquisition phase (Herwig et al., 2007; Herwig & Waszak, 2009). The absence of a congruency for the forced-choice subgroup led the authors to the conclusion that action control operates in two distinct control modes. They posit an intention-based control mode, as experienced in free-choice acquisition, and a stimulus-based control mode, as reflected in forced-choice acquisition (cf. Wenke, Waszak, & Haggard, 2009, and Keller et al., 2006, for further support of two distinct action control modes). According to their theory, both control modes imply the acquisition of associations, but these are of different types: While acting in an intention-based mode, action-effect associations as supposed by the ideomotor principle are acquired. In contrast, acting in a stimulus-based mode leads to the buildup of stimulus-response associations. They argue that in the test phase previously acquired action-effect (not stimulus-response) associations are used for action selection. Therefore, the forced-choice acquisition group, which has not adopted any action-effect associations according to this theory, cannot rely on any action-effect associations in the test phase. The assumption of two different action control modes was further substantiated by the absence of congruency effects for purely stimulus-based actions in several independent studies utilizing the action-effect acquisition paradigm (Herwig & Horstmann, 2011; Herwig & Waszak, 2012) or other paradigms derived from experiments on action-effect compatibility (Pfister et al., 2010; Pfister, Melcher, Kiesel, Dechent, & Gruber, 2014).

An alternative explanation for the absence of congruency effects in forced-choice acquisition groups was derived by Pfister et al. (2011). They proposed that participants in both control modes acquired action-effect associations, but only participants of the free-choice acquisition group actually made use of these associations during the forced-choice test phase. Their reasoning rested on two empirical findings. First, although using paradigms

different to the standard two-phase design, numerous studies on the acquisition of action-effect associations have provided evidence that action-effect learning is also present in forced-choice action mode as evidenced by experiments on action-effect compatibility (Kunde, 2001; Kunde, 2003; Kunde et al., 2004) and stimulus-effect compatibility (Hommel, 1993, 1996). Second, Pfister and colleagues (2011) conducted a study similar to the one of Herwig et al. (2007) with a free- and a forced-choice acquisition group (i.e., participants performed either free- or forced-choice key presses that contingently produced distinct effect tones), but changed the design of the test phase from forced-choice to free-choice. In contrast to Herwig et al. (2007), Pfister et al. (2011) observed a preference for acquisition-congruent responses, irrespective of the action mode of the acquisition phase. That is, both the free- and forced-choice acquisition group responded to the former effect tones more often with the key press that produced this tone during acquisition. As an overall conclusion, the authors suggested that action-effect learning occurs independent of the action control mode, while the action control mode determines whether existing action-effect associations are used to guide behavior or not.

The controversy regarding the impact of action control modes became even more pronounced, as Dignath et al. (2014) found reliable congruency effects in both free- and forced-choice test phases, when action-effect associations had been acquired in a forced-choice action mode (i.e., key presses to the left/right in response to a colored stimulus that resulted in distinct effect tones). This finding challenged both, the account of Herwig et al. (2007) that action-effect associations are not acquired in a stimulus-based action mode, and the hypothesis of Pfister et al. (2011) that action-effect associations acquired in a stimulus-based action mode are not used for behavioral control. To date, no clear consensus has therefore been reached with respect to the role of the action mode during acquisition in the manual domain.

---

## 5.2 ACQUISITION AND EFFECTS OF ACTION-EFFECT ASSOCIATIONS IN THE OCULOMOTOR DOMAIN

While certain manual actions, for example, finger pointing, are used to reach specific goals (e.g., conveying directions), also simple eye movements can be used with specific intentions in mind: Humans perform task-dependent eye movements to perceive the information which is relevant for them. This was shown in the context of scene perception studies (Yarbus, 1967), and later with respect to actions of daily life (Land, 2006; Land & Hayhoe, 2001; Land, Mennie, & Rusted, 1999). These studies have demonstrated that anticipation processes guided saccade paths, suggesting that eye movements can also be considered a goal-oriented action modality with the purpose of perceiving the post-saccadic object (Huestegge & Koch, 2009, 2010a). Consequently, research on intention-based action control, which was predominantly conducted within the manual action domain, has also started to focus more intensely on anticipation-based oculomotor control in the last few years (Herwig & Horstmann, 2011; Huestegge & Kreutzfeldt, 2012; Verschoor et al., 2013).

Among them, a pioneer study of Herwig and Horstmann (2011) is of special relevance for this work because of two reasons. First, the authors claimed that gaze plays a special role in the social (animate) environment, thereby emphasizing the importance to study oculomotor control processes in the *social domain*. Second, their study also directly contrasted the free-choice with the forced-choice action mode during acquisition, thereby contributing to the ongoing debate on the role of the action mode during acquisition, but now in the context of the oculomotor domain. In the following paragraphs, I will successively elaborate on these two points.

Regarding the special role of eye movements in the social domain, it appears evident that while eye movements cause merely perceptual effect signals in the inanimate environment, they are directed at other humans and elicit visible and gaze-contingent effects (e.g., a

smile on the face of the interaction partner) in the social environment. These thoughts are captured in the extension of the ideomotor theory to the social domain, the so-called theory of sociomotor action control (Kunde et al., 2018, see *Section 2.1.3*). It is based on the idea that social interaction builds on certain predictable regularities between an agent's action and the subsequent response of an interaction partner (e.g., a pointing gesture directs another person's orientation toward the respective location). The core message of the sociomotor principle is that similar processes as involved in ideomotor action control guide the selection and initiation of social actions. That is, agents acquire bi-directional action-effect associations between their own action and the response they invoke in others (i.e., social action effect), and anticipate this social action effect to initiate the motor action, which is associated with producing this effect.

To transfer the study of goal-oriented eye movements to the social domain, the study of Herwig and Horstmann (2011) used an adapted version of the action-effect acquisition paradigm: In an acquisition phase, participants performed saccades (to the left and right of the screen) toward neutral faces that contingently changed facial expression 100 ms after the saccade landed (e.g., the neutral face changed to happy for saccades to the face on the right, and to angry for saccades to the face on the left). In the subsequent test phase, the angry or happy face appeared centrally on the screen as an imperative stimulus. One group received the instruction to respond to the centrally presented face with a saccade toward the position where the facial expression was previously triggered (i.e., acquisition-congruent mapping) or to the other location (i.e., acquisition-incongruent mapping). The analysis of eye movements allowed for the combined testing of acquisition and application of action-effect associations within a single experimental setup. They observed shorter saccade latencies for the congruent action-effect mapping group, thereby demonstrating that action-effect associations were acquired, and replicating the well-established congruency effect from the manual

domain. To address the utilization of action-effect associations for action selection, they analyzed spatial saccade parameters during the acquisition phase, and observed that anticipations guided the initial saccade such that the mouth region was targeted more often for neutral faces changing to happy expression, and the eyebrow region was more often targeted for neutral faces changing into angry expression. Thus, their study revealed for the first time that humans acquire associations between saccades and their subsequent effects in the *animate* environment, and that the saccade-effect associations are used to guide eye movements.

Regarding the role of the action mode during acquisition, the study of Herwig and Horstmann (2011) was crucial for the current work by comparing the free-choice with the forced-choice action mode in the oculomotor domain. While participants in their study could freely decide for each trial where to look in the intention-based experiment (Herwig & Horstmann, 2011, Experiment 1), a white ellipse around one of the neutral faces indicated the direction in the stimulus-based experiment (Herwig & Horstmann, 2011, Experiment 2). In line with previous research from the Herwig group (Herwig et al., 2007; Herwig & Waszak, 2012), they only reported indicators of the acquisition and usage of action-effect associations for the intention-based version.

The role of action anticipation in oculomotor control in *inanimate* environments was already addressed in Experiment 1 of the current work (that was done in cooperation with Herwig and Horstmann). In the respective experiment, participants executed left versus right free-choice saccades toward traffic lights that triggered an action-contingent light switch in the upper versus lower part of the traffic light. Anticipation processes were reflected in spatial saccade shifts, that is, participants directed their initial saccade toward the location of the upcoming light shift. Together with the results of Herwig and Horstmann (2011), one can conclude that spatial saccade characteristics reflect anticipation in both animate and inanimate environments (see also Pfeuffer et al., 2016, for similar results in the manual domain). Further evidence of effect-based oculomotor control in the inanimate environment comes

from Huestegge and Kreutzfeldt (2012). The following experiment of this work was conducted to integrate and extend the abovementioned results with respect to mechanisms of oculomotor action control in social versus non-social contexts and with respect to the role of a stimulus- versus intention-based action mode.

### 5.3 EXPERIMENT 9

As outlined above, there are studies from both the manual (Pfister et al., 2011) and the oculomotor (Herwig & Horstmann, 2011) domain that addressed the role of action modes for the acquisition of action-effect associations. Over the past few years, researchers have been showing a growing interest for action control in social contexts (e.g., Müller, 2016; Pfister, Obhi, Rieger, & Wenke, 2014; Weller et al., 2019), thereby highlighting commonalities and differences to non-social contexts (Kunde et al., 2018). Thus, there are studies from both the manual and the oculomotor action domain that focus on the acquisition of action effects in either social (Herwig & Horstmann, 2011; Sato & Itakura, 2013) or non-social environments (Elsner & Hommel, 2001; Huestegge & Kreutzfeldt, 2012). What is still pending is a systematic comparison of an intention- versus stimulus-based way of acquiring action-effect associations (in an action-effect acquisition paradigm) using social and non-social action effects within the identical experimental design. Hence, the subsequent experiment was designed to bridge this gap.

The experiment relied on the action-effect acquisition paradigm from Elsner and Hommel (2001), and either assigned participants to a social or a non-social context group (using the schematic face or the abstract geometrical stimuli from Experiments 2–6 as stimuli). Then, a free- and a forced-choice manipulation was implemented in both contexts. All participants executed upward and downward eye movements toward on-screen targets,

---

which resulted in distinct visual effects (i.e., action effects). The effect stimuli of the acquisition phase were presented as imperative stimuli in the subsequent test phase together with a congruency manipulation to check whether participants acquired bi-directional associations between their actions and the resulting effects. Spatial saccade characteristics during acquisition were analyzed to assess the usage of action effects for saccade planning. In line with observations by the Herwig lab (e.g., see Herwig & Horstmann, 2011), I predicted congruency effects to be present throughout all conditions, but they should be attenuated in forced- (vs. free-)choice conditions. Further, I expected action-effect anticipation in oculomotor control to be reflected in corresponding spatial saccade characteristics during acquisition, similar to corresponding observations observed in social (Herwig & Horstmann, 2011) and non-social contexts (Experiment 1, see *Section 3.1*). Additionally, I assessed whether action control differed between social and non-social context conditions. However, given the results from the experiments reported so far within this work, I did not expect to observe meaningful differences between context conditions.

## 5.3.1 METHODS

### 5.3.1.1 Participants

Based on previous results using the present experimental design in the oculomotor domain (Herwig & Horstmann, 2011), I conducted a power analysis using G\*Power 3 (Faul et al., 2007) which suggested to test 20 participants (ten per congruency group) in each of the four parts of the experiment (social vs. non-social context, free- vs. forced-choice) to observe a congruency effect with considerable probability (power = .95,  $\alpha$  = .05). In total, 88 participants were recruited, but five participants had to be excluded due to unusually high error rates (> 30% in either of the cells of the experiment design), and another three participants due to a low number of valid data (less than two-third of the recorded data) in at least one cell of the design. Data of the remaining 80 participants were included in the analysis

(mean age: 23.7 years,  $SD = 4.1$  years, age range = 18 – 39 years, 25 male). Participants received either course credits or payment and reported normal or corrected-to-normal vision. They were naïve with respect to the purpose of the experiment and gave informed consent before completing the study.

### 5.3.1.2 Stimuli

The same schematic face stimuli as used in Experiment 3 served as stimuli for the social context condition, and the same abstract geometric stimuli as used in Experiment 5 were used in the non-social context condition (see *Figure 8*, panels B and D). Stimuli were presented in vertical ellipses with a size of  $3.5^\circ \times 5.3^\circ$  of visual angle (maximum width  $\times$  maximum height) in the acquisition phase and of  $2.6^\circ \times 4.0^\circ$  in the test phase. The identical acoustic stimuli as in Experiments 2–6 were used.

### 5.3.1.3 Procedure

Following the experimental setting implemented by Herwig and Horstmann (2011), who studied facial emotional changes instead of gaze changes, the experiment consisted of an acquisition phase (in either free- or forced-choice version) and a subsequent test phase (see *Figure 15*). Before the experiment started, participants were familiarized with the acoustic stimuli used during the experiment. Each trial of the acquisition phase started with the concurrent presentation of a central fixation cross ( $0.4^\circ \times 0.4^\circ$  of visual angle) and two baseline stimuli (i.e., schematic face stimuli with closed eyes or abstract stimuli without a black square, respectively). Baseline stimuli served as saccade targets and were displayed at  $7.9^\circ$  of visual angle above and below the fixation cross. I used a flicker manipulation to focus the participant's attention on the fixation cross during the fixation interval, that is, after a variable interval of 300 to 550 ms the fixation cross was suspended for a duration of 200 ms. The total duration of the fixation interval varied between 1,000–1,500 ms.



---

Participants were instructed to direct their gaze from central fixation toward one of the peripheral saccade targets. In the free-choice condition, the low-pitched tone served as starting signal, and the choice of saccade direction (up/down) was up to the participants. They were told to choose each direction (up/down) about equally often and to avoid any possible regularities in saccade direction sequences. After each block, on-screen feedback regarding the number of up/down saccades was provided to further facilitate compliance with the instructions. In the forced-choice condition, participants were instructed to shift their gaze from central fixation upward or downward toward the target regions in accordance with the stimulus pitch (high- or low-pitched tone). Participants in both choice conditions (free- vs. forced-choice) received the instruction to execute their response as quickly (and in the forced-choice condition additionally as accurately) as possible, and return to the fixation cross immediately afterwards.

In both choice conditions, each saccade triggered a particular effect within the saccade target, like in Experiments 2–6. Contingent upon the participant's gaze shift toward the upper or lower target area, the stimulus in the target area changed 200 ms after the saccade landed within the target area. Crucially, saccade direction determined which change occurred as an effect of the saccade. For example, for one half of the participants, a saccade toward the upper face consistently triggered the onset of a direct gaze, and a saccade toward the lower face triggered the onset of an averted gaze (or the corresponding equivalents for abstract stimuli). The opposite assignment was provided for the other half of participants. The direction of the averted gaze or lateral square stimulus (i.e., averted gaze/lateral square to the left/right) was counterbalanced. That is, one half of the participants consistently experienced a left gaze/left square when looking at the averted gaze/lateral square target, while the other half of participants consistently experienced a right gaze/right square when looking at the averted gaze/lateral square target. The effect stimulus was presented for 500 ms, and the next trial started after a blank interval (blank screen) of 200 ms. Like in Experiments 2–6, no

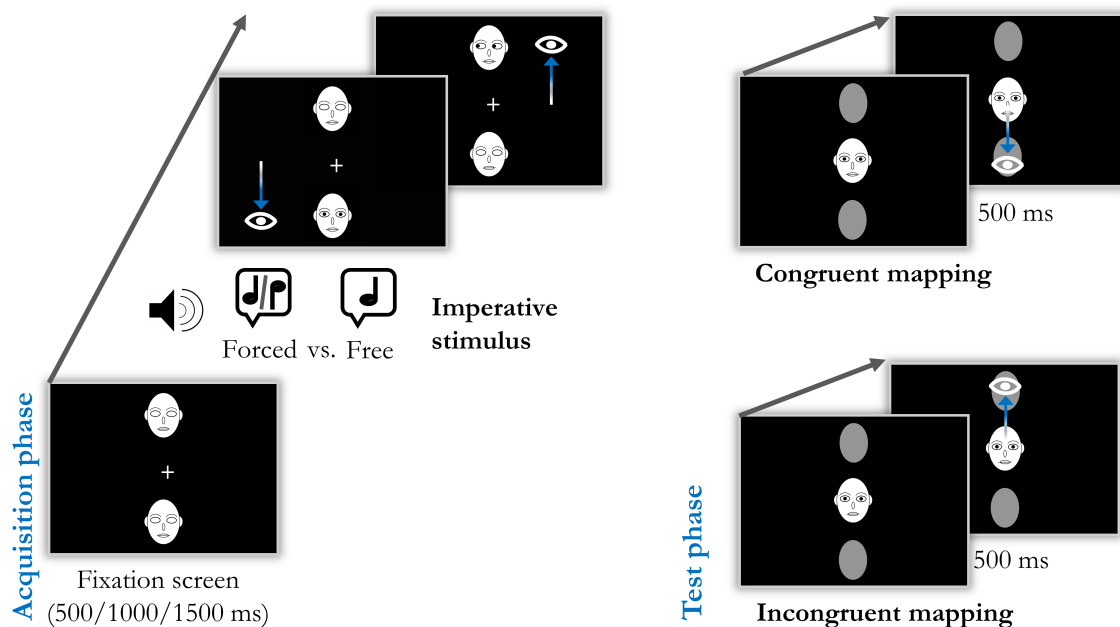
---

error feedback was provided in case of an erroneous saccade response in the forced-choice condition, and a saccade toward the upper or lower position *always* triggered the associated effect stimulus irrespective of whether the executed saccade response was in line with the response indicated by the pitch of the tone or not. If no saccade toward one of the targets had occurred within 1,500 ms after tone onset, the feedback message “*Keine Reaktion*” (German for “no response”) appeared for a duration of 300 ms and the next trial started after a blank interval (black screen) of 200 ms. Trial order was pseudo-randomized to prevent order effects. Thus, the number of switches between high-pitched and low-pitched auditory stimuli was roughly equal, and the maximum number of consecutive repetitions of the same auditory stimulus was restricted to six repetitions.

In the subsequent test phase, participants were instructed to respond with upward or downward saccades to the effect stimuli from the preceding acquisition phase. Crucially, participants were divided into two groups receiving different effect-action assignments. Half of the participants received an acquisition-congruent assignment, requiring them to respond to the presented effect stimulus with a saccade into the direction that triggered this effect in the preceding acquisition phase. For the other half of participants, an acquisition-incongruent assignment required to respond to the effect stimulus with a saccade into the direction opposite to that this effect has been associated with during acquisition. Note that the context condition (schematic face vs. abstract geometric stimuli) was consistent between acquisition and test phase for each participant. Each trial of the test phase started with the presentation of a fixation cross (for a variable duration of 1,000–1,500 ms). After that, one of the former effect stimuli was presented centrally on the screen, serving as an imperative stimulus for the participant. The effect stimulus was surrounded by two solid grey ellipses ( $2.6^\circ \times 4.0^\circ$ ) at a horizontal distance of  $7.9^\circ$  of visual angle, serving as potential saccade targets. The next trial started 1000 ms after a saccade toward one of the target areas was detected. If no such saccade had occurred within 1,500 ms, the next trial started immediately after presenting the

feedback message “*Keine Reaktion*” (German for “no response”) for a duration of 500 ms. The trial order was pseudo-randomized, with the maximum number of consecutive repetitions of the same effect stimulus restricted to five repetitions.

Each experimental session comprising acquisition and test phase lasted around 30 minutes. The acquisition phase comprised seven experimental blocks with 32 trials each (224 trials in total), the test phase consisted of four experimental blocks with 24 trials each (96 trials in total). A calibration of the eye tracker was performed prior to each experimental block.



*Figure 15.* Setup of Experiment 9 (for schematic face stimuli). In the acquisition phase, an auditory stimulus (high- or low-pitched tone for forced-choice; only low-pitched tone for free-choice) signaled the participant to perform a corresponding gaze shift upward or downward (forced-choice) or to choose between performing a gaze shift upward or downward (free-choice) toward one of the peripheral faces. As an effect of the saccade, either a direct or averted gaze was displayed (contingent upon the targeted face) for a duration of 500 ms, before the next trial started. In the subsequent test phase, the former effect stimuli were presented as imperative stimuli requiring gaze shifts (up/down) according to a specific assignment: Half of the participants received an acquisition-congruent assignment (top right), while the other half received an acquisition-incongruent assignment (bottom right).

### 5.3.1.4 Design and analysis

The experiment involved the between-subjects factors *congruency* (acquisition-congruent vs. acquisition-incongruent mapping), *choice condition* (free- vs. forced-choice), and *context* (social vs. non-social). Furthermore, the within-subjects factor *direction* of the action effect (direct vs. averted gaze in the social context; central vs. lateral effect in the non-social context) was manipulated. For the acquisition phase, I analyzed the direction of saccade responses (up vs. down) as well as the relative angle deviation (rAD) as spatial measures. The rAD is the absolute amount of the mean deviation (in degrees of visual angle) of the response saccade from a perfectly horizontal saccade as measured from the starting point of the response saccade. For the averted gaze or the lateral effect direction condition, respectively, positive rAD values were indicators of spatial effect anticipation, irrespective of the orientation of the averted gaze/lateral effect direction to the left or right (see Herwig & Horstmann, 2011, and Experiment 1 of this work, *Section 3.2*, for a similar approach)<sup>15</sup>. In the test phase, saccade latency and error rates were analyzed as performance measures.

### 5.3.1.5 Preprocessing of saccade data

For the analysis of the saccade data during acquisition, different criteria were applied to detect the response saccade toward the upper or lower saccade target in the free- versus forced-choice condition. Originally, I constructed both the free- and forced-choice conditions in a maximally comparable way to apply the identical criterion for the determination of the response saccade. Analogous to Experiments 2–6, the saccade after onset of the imperative auditory stimulus with an amplitude of at least one third of the distance between the fixation cross and the center of the saccade target was defined as the response saccade. However, a considerably large amount of participants of the free-choice condition did not follow

---

<sup>15</sup> Please note, however, that for the action effects in the face stimuli, it is quite unlikely to expect any measurable effects on rAD, since the location of the action effect for the averted gaze condition is associated with a relatively minor change occurring in both halves of the face stimulus, and not restricted to a well-defined area within the stimulus (see *Section 5.3.3* for more details). Nevertheless, I still decided to analyze this parameter.

---

the given instructions correctly and initiated their saccades toward the upper or lower saccade target constantly *before* the auditory signal was presented, which was meant to serve as the starting signal. Thus, I applied the original criterion to identify the response saccade to the forced-choice condition only. For the free-choice condition, however, I used the flicker manipulation occurrence as a reference time point instead of the audio signal (see *Section 5.3.1.3* for procedural details) while sticking to the amplitude-based saccade detection criterion. In addition to that, only response saccades that started within the predefined interest area around the fixation cross were analyzed in the free-choice condition, which served as an additional intervention to ensure that the actual response saccade was detected despite the temporal shift of the time window. To ascertain whether the latter method results in reliable data, the data of the forced-choice condition were analyzed using both saccade detection criteria, and I obtained identical result patterns for both ways of analyzing the rAD. Another eight subjects of the free-choice condition had to be excluded from analysis of the acquisition data, since the use of the adapted criterion returned low numbers of valid data (less than two-third of the recorded data) in at least one cell of the design for these subjects. While the use of different detection criteria is unproblematic with respect to the analysis of the spatial measure (rAD), I refrained from analyzing saccade latencies for the acquisition phase, since different points in time were used as starting points for the detection of the saccade responses in the free- versus forced-choice conditions. However, these circumstances do in no way limit the significance of the overall results, since the relevant measure for the present purpose were RTs during the *test* phase. In the test phase, which was of forced-choice type for all subjects, participants had to wait for the appearance of one of the former effect stimuli as imperative stimulus to initiate a response saccade into the direction indicated by the respective effect-action assignment.

## 5.3.2 RESULTS

### 5.3.2.1 Acquisition phase

Prior to analysis, 8.4% of the total number of trials (32 participants with 224 trials each resulting in 7168 trials in total) of the free-choice condition and 2.1% (40 participants with 224 trials each resulting in 8960 trials in total) of the forced-choice condition were removed after applying the cleansing procedure as described for Experiments 2–6.

In the free-choice acquisition phase, the frequencies of upward and downward response saccades was close to the instructed balanced distribution. The average proportion of upward versus downward saccades of all free-choice trials amounted to 50.5% versus 49.5%, respectively. The analysis of the rADs was based on correct trials only. A three-way ANOVA with the group factors context and choice condition and the within-subject factor direction of the action effect was conducted for rADs. The significant interaction of direction and context,  $F(1,68) = 12.37, p = .001, \eta^2_p = .15$ , demonstrated greater rADs for lateral ( $M = 2.59^\circ, SE = 0.42^\circ$ ) versus central effect direction ( $M = 0.25^\circ, SE = 0.39^\circ$ ) in the non-social context condition,  $p < .001$ , while rADs did not differ in the social context condition for averted ( $M = 0.93^\circ, SE = 0.41^\circ$ ) versus direct gaze direction ( $M = 0.94^\circ, SE = 0.38^\circ$ ),  $p = .974$  (see *Figure 16*). The effect of direction was significant,  $F(1,68) = 12.05, p = .001, \eta^2_p = .15$ , while the effect of context was not significant,  $F(1,68) = 1.13, p = .291, \eta^2_p = .02$ . All other effects were far from the level of significance, all  $F_s < 1$ .

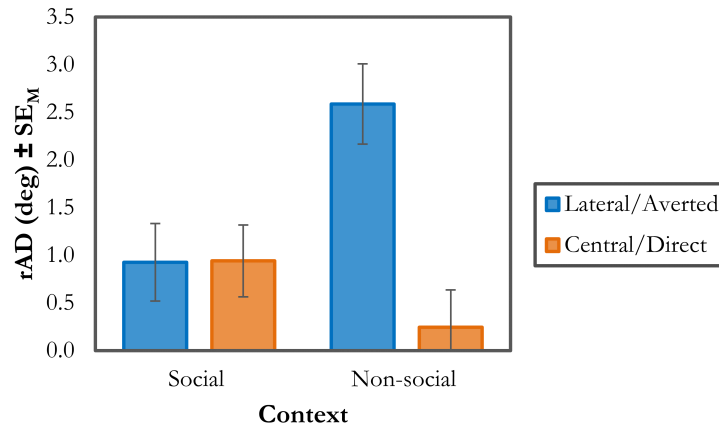


Figure 16. Relative angle deviation (rAD in °) in the acquisition phase of Experiment 9 as a function of direction (direct gaze/central effect vs. averted gaze/lateral effect) and context (social vs. non-social). Error bars represent standard errors of the mean ( $SE$  in °).

### 5.3.2.2 Test phase

For the test phase data, I applied the same cleansing procedure as in Experiments 2–6. Thus, prior to analysis, 1.7% out of the total number of trials (80 participants with 96 trials each resulting in 7,680 trials in total) were removed. The original design included the within-subjects factor direction (direct vs. averted gaze direction in the social context; central vs. lateral effect direction in the non-social context condition). Since none of the statistical comparisons including this factor reached significance in the test phase, I decided to average data across this factor. Saccade latencies of correct responses were submitted to a three-way ANOVA with the group factors congruency, choice condition, and context. In summary, there was no effect of congruency as well as of choice condition, both  $F_s < 1$ . None of the interactions reached significance, all  $F_s < 1$  (except for the interaction of choice condition and congruency,  $F(1,72) = 1.38, p = .244, \eta^2_p = .02$ ). The main effect of context was the only significant effect,  $F(1,72) = 12.63, p = .001, \eta^2_p = .15$ . As can be seen in Figure 17, saccade latencies were shorter by 55 ms in the non-social context as compared to the social context ( $M = 387$  ms,  $SE = 11$  ms vs.  $M = 442$  ms,  $SE = 11$  ms).

The overall error rate amounted to 7.6% ( $SE = 0.7\%$ ). The effect of context on error rates was not significant,  $F(1,72) = 3.13$ ,  $p = .081$ ,  $\eta^2_p = .04$ . None of the remaining effects revealed significant results, all  $F$ 's  $< 1$ .

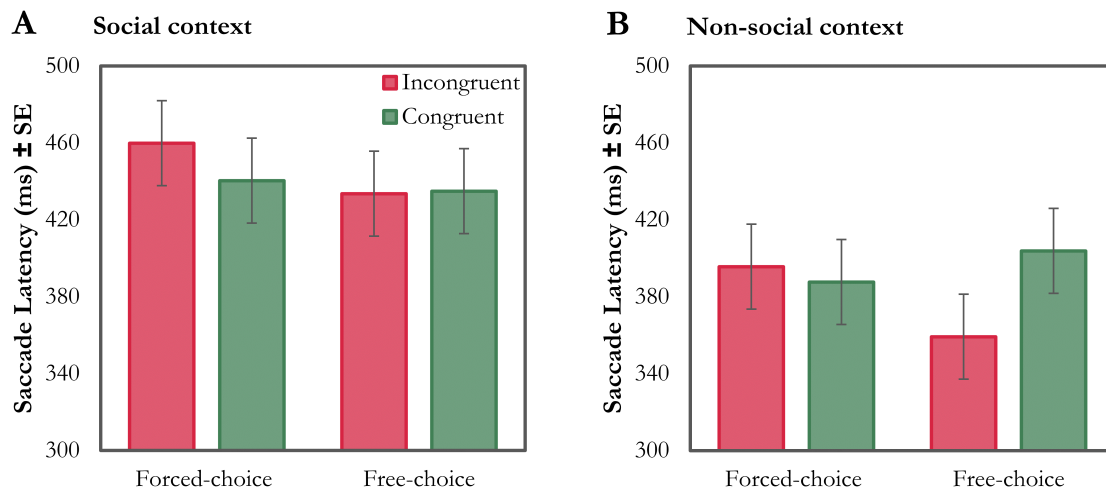


Figure 17. Saccade latencies (in ms) of the test phase of Experiment 9 as a function of congruency (congruent vs. incongruent), choice condition (forced- vs. free-choice), and context (social, Panel A vs. non-social, Panel B). Error bars represent standard errors of the mean ( $SE$  in ms).

### 5.3.2.3 Exploratory analysis (test phase)

To identify potential mechanisms behind the absence of any congruency effects in the test phase, a median split was performed by separating data into fast and slow RTs for each participant and condition separately. Response time data and error rates were subsequently submitted to a four-way repeated-measures ANOVA with speed (fast vs. slow) as a within-subjects factor, and congruency, choice condition, and context as group factors. Note that I only report effects involving the relevant factor speed here, as the result pattern for non-speed related effects is identical with what has been reported above for the three-way ANOVAs.

When analyzing RTs, none of the interactions involving the factor speed reached significance, all  $F$ 's  $< 1$ . The (self-explanatory) effect of speed was the only significant effect,  $F(1, 72) = 643.30$ ,  $p < .001$ ,  $\eta^2_p = .90$ .



The analysis of error rates revealed a significant effect of speed,  $F(1, 72) = 10.11, p = .002, \eta^2_p = .12$ , demonstrating that participants committed more errors in the context of fast responses ( $M = 8.7\%$ ,  $SE = 0.8\%$ ) as compared to slow responses ( $M = 6.6\%$ ,  $SE = 0.7\%$ ). None of the interactions including speed and congruency reached significance, all  $F_s < 1$ . The interaction of speed and choice condition was significant,  $F(1, 72) = 4.76, p = .032, \eta^2_p = .06$ , indicating that error rates for fast versus slow RTs only differed when participants were in the free-choice group. Also, the three-way interaction of speed, choice condition, and context was significant,  $F(1, 72) = 6.62, p = .012, \eta^2_p = .08$ : While there was a significant interaction of speed and choice condition in the non-social context (as described above), no such interaction occurred in the social context condition.

Moreover, I conducted a separate three-way ANOVA with the group factors congruency, choice condition, and context for RTs involving the trials from the first block only (initial performance analysis). The results reflect the pattern of the overall analysis: The effect of context revealed significant results,  $F(1, 71) = 5.56, p = .021, \eta^2_p = .07$ , showing the same pattern as the analysis involving all blocks. None of the remaining effects was significant, all  $F_s < 1$ .

### 5.3.3 DISCUSSION

Experiment 9 was conducted to compare two ways in which an oculomotor action effect is specified, namely in an intention-based (free-choice) and a stimulus-based (forced-choice) mode when participants act in a social or non-social context, respectively. To start with the results of the test phase, I did neither observe any influence of congruency on RTs nor error rates, thereby failing to replicate the congruency effect as an indicator of the acquisition of bi-directional action-effect associations (Elsner & Hommel, 2001; Herwig & Horstmann, 2011). Thus, oculomotor action-effect learning was not evident using the well-established two-phase action-effect acquisition paradigm. Performance during test phase was not

affected by the free- versus forced-choice mode. Only the context variable affected performance in the test phase such that participants responded faster (but tended to commit more errors) when presented with abstract non-social (vs. social face) stimuli. Exploratory analyses further corroborated the absence of any congruency influence on performance when fast versus slow reactions were analyzed separately. Overall, the results of the test phase performance did not demonstrate any anticipation effects, independent of the particular context or choice conditions.

In contrast, the analysis of spatial saccade parameters during acquisition revealed that participants anticipated the action effect and directed their response saccade into the direction of the anticipated effect, but only in a non-social context. That is, when abstract geometric stimuli served as action effects, the angle of the response saccade deviated significantly more from a vertical saccade (up/down) for the lateral (vs. central) effect direction, indicating an influence of the upcoming effect on saccade planning. When schematic faces served as action effects, no such anticipation effects were reflected in spatial saccade characteristics. One might argue that the absence of such a spatial oculomotor marker of effect anticipation is due to the social context per se – however, the demonstration of such markers with face photographs as stimuli (Herwig & Horstmann, 2011) likely disapproves this hypothesis. Most likely, the absence might be due to the stimulus material per se. I observed spatial oculomotor markers of effect anticipation with abstract stimuli, where the gaze-contingent change in the object was especially pronounced. In contrast, the effect location for the averted gaze condition is not restricted to a well-defined area of the stimulus, but rather associated with a relatively minor change occurring in both halves of the face stimulus (i.e., the eyeball shifts to side in both the left and right eye). Therefore, the lack of anticipation effects in the spatial oculomotor parameters for the face stimuli is not surprising and does clearly not rule out that effect anticipation also occurred in a social context.

---

Another important purpose of this experiment was to compare an intention- and a stimulus-based way of acquiring action-effect associations in an action-effect acquisition paradigm. Due to the absence of any congruency effects in the test phase performance as indicators of bi-directional action-effect learning, I am not able to assess any differences in action-effect learning as reflected in congruency effects depending on the action mode. However, the spatial anticipation effects observed in both, an intention-based and a stimulus-based mode during acquisition, suggest that action-effect learning is in fact independent of the action mode during acquisition. In view of these results, I agree with the approach as summarized in Pfister, Melcher et al. (2014), which represents a synthesis of the different lines of arguments (see *Section 5.1*). Specifically, they argued that the tendency that effect-based action control is more pronounced for an intention-based action mode compared to a stimulus-based action mode during acquisition (Herwig et al., 2007) seems to be restricted to situations with variable and context-dependent action-effect relations (Pfister et al., 2010). In contrast, if action-effect associations are stable, acquired action-effect associations are used for action selection independent of an intention- versus stimulus-based action mode (Pfister et al., 2011; Pfister & Kunde, 2013; Wolfensteller & Ruge, 2011).

Regarding potential reasons why oculomotor action-effect learning was not evident in this experiment, one might speculate that either action-effect learning between saccades and the following visual effects did not take place at all, or that learning did occur but failed to surface in the test phase (Pfister et al., 2011; Pfister, 2019). In light of effects of spatial anticipation as demonstrated in the experiment above (even though for a non-social context only; see the previous paragraph for details), I would rather assume that action-effect links have been learned, but the reason for the absence of any test phase effects is to be found in the paradigm itself. While several studies from the manual domain are relying on action-effect acquisition paradigms (Beckers et al., 2002; Elsner & Hommel, 2001; Elsner & Hommel, 2004; Pfister et al., 2011) yielded converging evidence for action-effect learning, a single

study from the oculomotor domain investigated oculomotor action-effect learning with the help of this two-phase action-effect acquisition paradigm (Herwig & Horstmann, 2011). Herwig and Horstmann report two experiments from which only one (see their Experiment 1) revealed evidence for the learning of action-effect associations. This divergence of results was explained by different action modes during acquisition. Furthermore, if you assume that the results of their Experiment 1 might represent a statistical Type-I error, there is no other study present in the literature that demonstrated *oculomotor* action-effect learning within the action-effect acquisition paradigm. Hence, the present failure to demonstrate oculomotor action-effect learning might not be seen as evidence for the absence of action-effect learning, but might rather point toward a reduced sensitivity of the two-phase action-effect acquisition paradigm to reveal action-effect learning, especially in eye movements. At this point, it is important to note that the experimental procedure of Experiment 9 of this work resembles the two-phase action-effect acquisition paradigm as implemented by Herwig and Horstmann (2011). However, these two studies also differ in some aspects, for example in the type of action effects (gaze responses of schematic face stimuli in Experiment 9 vs. emotional face responses of face photographs in their study) and in the direction of the to-be-executed response saccade (upward/downward in Experiment 9 vs. to the left/right in their study). Together, these procedural differences could also account for the observed diverging results.

When assuming that the lacking evidence in favor of action-effect learning originates from the chosen experimental design in the current work, a critical examination of the chosen congruency manipulation in the test phase is of utmost importance. To be precise, the effect stimuli from the preceding acquisition phase were presented as imperative stimuli in the test phase, which should prime the associated action, according to the common ideomotor principle (cf. Elsner & Hommel, 2001). If the required action is congruent to the primed action, RTs should be facilitated. The chosen test phase design of Experiment 9 of the present work required participants to attend and process the previous effect stimuli (see

---

also Beckers et al., 2002; Hoffmann et al., 2009), thereby increasing the chance for the imperative effect stimulus to retrieve the associated action. There are other paradigms, in which neutral targets served as imperative stimuli. These were preceded by a short presentation of one of the effect stimuli from acquisition, which primed or interfered with the requested response, thereby manipulating congruency within subjects (e.g., Sato & Itakura, 2013). In such experimental designs, it is in principle possible to assume that participants ignore the prime, which might eventually preclude congruency effects to emerge. According to this logic, however, one would actually have expected absent congruency effects in the previous experiments of this work (namely, Experiments 2-8), since participants did not necessarily had to attend to the task-irrelevant prime stimulus in these experiments. In light of this methodical difference between the two designs, the absence of any evidence in favor of action-effect learning in Experiment 9, where participants were required to attend to and to process the effect prime, is even more astonishing.

This leads me to the question if the two-phase action-effect acquisition paradigm is only suitable to reveal manual, but not oculomotor action-effect learning due to differences between these action modalities. Due to the strong evidence in favor of action-effect learning as revealed by the task-irrelevant prime-paradigm (Experiments 3, 5, 7, and 8), I propose the following explanation. Eye movements might be more flexible and adaptive than manual actions due to the constant requirement to adapt to changing visual input for humans. That is, the action-effect associations acquired during acquisition dissipate and adapt to the context of the test phase rather quickly for eye movements such that priming processes are active during the very first trials of the test phase only, and prevent global, measurable congruency effects to emerge. Since no congruency effects were observed for the initial performance analysis (including data from the first block only), it can be concluded that the proposed adaptation mechanism is likely to set in very quickly, that is, during the very first trials of the test phase already. When the action-effect associations are re-instantiated in every trial, as it

is the case within the task-irrelevant prime paradigm, the task-irrelevant stimulus has the power to prime the associated response, which is reflected in stable congruency effects. If this would be a valid assumption, repeating Experiment 9 with manual instead of oculomotor actions should enable me to observe congruency effects. Thus, to test the assumption that the two-phase action-effect acquisition paradigm has a reduced sensitivity for revealing action-effect learning in eye movements due to a less long-lasting memory (or binding) of action-effect associations, I conducted the following experiment, which represents a conceptual replication of Experiment 9, but transferred to the manual domain.

## 5.4 EXPERIMENT 10

The results of Experiment 9 clearly showed that the action-effect associations as experienced during an acquisition phase did not affect performance in terms of saccadic RTs and error rates in a subsequent test phase, and that this pattern was independent of whether participants acted in an intention-based or stimulus-base mode during acquisition. Experiment 10 was conducted to test whether the absence of any congruency effects in Experiment 9 can be attributed to the use of eye movements as an action modality. Therefore, Experiment 9 was repeated in the manual domain. That is, the general procedure of Experiment 10 was identical to the forced-choice procedure of Experiment 9 with the only difference being that participants did not use their gaze, but manual key presses as response device. Since the focus of Experiment 10 was no longer on the role of the action mode during acquisition, I refrained from implementing a free-choice acquisition phase given that Dignath et al. (2014, Experiments 2A and B) observed greater effect sizes when acquisition was conducted in a forced- (vs. free-)choice mode. Given numerous studies from the manual domain demonstrating evidence for the acquisition of action-effect associations within action-effect acquisition paradigms (e.g., Elsner & Hommel, 2001; Elsner & Hommel, 2004; Pfister et al., 2011),

I expected that key press actions and their visual effects should become associated during acquisition, which in turn affects key press latencies in response to the former effect stimuli in a subsequent test phase. Thus, RTs should be faster for an acquisition-congruent mapping as compared to an acquisition-incongruent mapping. As in Experiment 9, participants acted either in a social or a non-social context to assess differences and commonalities underlying social versus non-social (manual) action control.

## 5.4.1 METHODS

### 5.4.1.1 Participants

Based on observed effect sizes in the study of Dignath et al. (2014) from the manual domain, the power analysis using G\*Power 3 (Faul et al., 2007) indicated a sample size of 48 (24 per each congruency group) as sufficient to observe a congruency effect with considerable probability (power = .80,  $\alpha$  = .05). Thus, forty-eight participants were recruited. One participant had to be excluded because performance in the acquisition phase was obviously executed in opposition to the instructions (resulting in an error rate of close to 100%). Data of another two participants were excluded due to unusually high error rates (> 30%) in one of the experimental cells in the test phase. The data of the remaining 45 participants (mean age: 23.5 years,  $SD$  = 4.3 years, age range = 18 – 43 years, 10 male) were analyzed. Participants received either course credits or payment and reported normal or corrected-to-normal vision. They gave their informed consent before completing the study and were naïve with respect to the purpose of the experiment.

### 5.4.1.2 Stimuli and procedure

The set of stimuli used in Experiment 10 was identical to the one used in Experiment 9. In general, the procedure was identical to Experiment 9, but instead of instructing participants to saccade toward the upper or lower target according to the pitch of the tone,

participants were asked to respond to the high- versus low-pitched tones with spatially corresponding key presses on a standard QWERTZ computer keyboard, that is, with up or down arrow key presses. Identical to the procedure of Experiment 9, the baseline target changed to a schematic face with direct or averted gaze direction (or the corresponding equivalents for abstract stimuli) in response to their key press actions. Each experimental session comprising acquisition and test phase lasted around 35 minutes.

#### 5.4.1.3 Design and analysis

Prior to analysis, all trials with response anticipations (RTs < 100 ms) were removed. For the analysis of the acquisition phase, mean RTs and error rates are reported. For the analysis of the test phase data, RTs of correct responses and error rates were submitted to separate ANOVAs with the group factors *context* (social vs. non-social) and *congruency* (acquisition-congruent vs. acquisition-incongruent) and the within-subjects factor *direction* of the action effect (direct gaze vs. averted gaze in the social context; central effect vs. lateral effect in the non-social context).

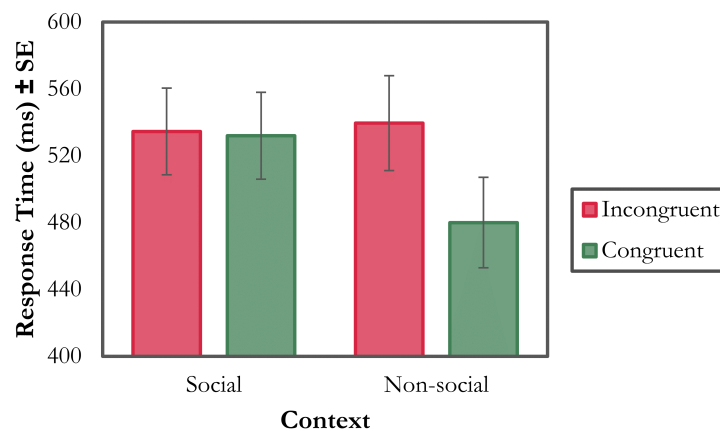
### 5.4.2 RESULTS

Prior to analysis, 0.4% of all acquisition phase trials were removed, which corresponded to 38 of the total number of trials (45 participants with 224 trials each resulting in 10,080 trials in total). Mean RT and mean error rate during acquisition amounted to 416 ms ( $SE = 13$  ms) and 2.6% ( $SE = 0.4\%$ ), respectively. To analyze the test phase data, I removed 0.5% of all trials, which was equivalent to 21 of the total number of trials (45 participants with 96 trials each resulting in 4,320 trials in total). The original design intended to include the factor *direction* within the analysis, but since none of the statistical effects involving this factor was significant for neither RTs nor error rates (all  $ps > .05$ ) the factor *direction* was removed from analysis (cf. Experiment 9 for the same procedure). RTs amounted to 506 ms ( $SE = 19$  ms) versus 537 ms ( $SE = 19$  ms) for the acquisition-congruent versus acquisition-



incongruent group (see *Figure 18*). The two-way ANOVA involving the factors congruency and context revealed no significant effects, neither for congruency,  $F(1, 41) = 1.33, p = .255, \eta^2_p = .03$ , nor for context,  $F < 1$ , nor for the interaction of both,  $F < 1$ .

The error rate amounted to 2.6% ( $SE = 0.5\%$ ) versus 3.5% ( $SE = 0.5\%$ ) for the acquisition-congruent as compared to the acquisition-incongruent condition. As for the analysis of RTs, the ANOVA for the error rates did not return any significant results, all  $F_s < 1$  (except for the main effect of congruency,  $F(1, 41) = 1.23, p = .274, \eta^2_p = .03$ ).



*Figure 18.* Mean RTs (in ms) in the test phase of Experiment 10 depicted as a function of congruency (congruent vs. incongruent) and context (social vs. non-social). Error bars represent standard errors of the mean ( $SE$  in ms).

### 5.4.3 DISCUSSION

Experiment 10 aimed to (re-)apply the two-phase action-effect acquisition paradigm to the manual domain while keeping the entire setup as close as possible to the setup of Experiment 9, which was conducted in the oculomotor domain. Observing congruency effects in Experiment 10, which only implemented forced-choice actions due to their greater potential to reveal congruency effects (Dignath et al., 2014), would indicate that the lack of evidence for the acquisition of action-effect associations in Experiment 9 is due to the action modality: While participants responded with eye movements in Experiment 9, they per-

formed key presses in Experiment 10. However, the results demonstrate no systematic differences in neither RTs nor error rates depending on whether participants were in the acquisition-congruent or acquisition-incongruent group. In addition, no influence of a social versus non-social context was evident. In general, these findings replicate the results of Experiment 9, but with manual actions. The results of both experiments differ only in the fact that saccades were initiated faster in a non-social (vs. social) context, while there was no such difference evident for key presses. These results, together with the results of Experiment 10, suggest an insensitivity of the present paradigm to reveal action-effect learning, not only for eye movements, but also for key presses, and have important implications, which will be discussed in the following section.

## 5.5 CONCLUSION OF CHAPTER 5

Both experiments presented in this chapter failed to demonstrate empirical markers of oculomotor (Experiment 9) or manual (Experiment 10) action-effect learning in terms of congruency effects in RTs using the two-phase action-effect acquisition paradigm. Participants were responding with about the same speed to initiate a saccade or a key press in response to an imperative stimulus, independent of whether the required action matched the action that was associated with the imperative stimulus during acquisition or not. Since I did neither observe any congruency effects in Experiment 9 nor Experiment 10 (except for the congruency effect in a spatial saccade parameter during acquisition in Experiment 9), the main research question of the present chapter, that is, to assess the impact of different action modes during acquisition on action-effect learning, recedes into the background. Instead, I will discuss the suitability of the paradigm itself for the investigation of ideomotor manual and oculomotor action control.

---

The absence of any congruency effects in Experiments 9 and 10 is surprising with respect to several aspects. First, there is notable empirical support for effect-based oculomotor control from experiments presented within this work (cf. Experiments 1, 3, 5, 7, and 8) and in previous literature (e.g., Herwig & Schneider, 2014; Herwig, Weiss, & Schneider, 2015; Huestegge & Kreutzfeldt, 2012; Verschoor et al., 2013; see Herwig, 2015 for a review). While most of these previously published studies relied on empirical paradigms other than the paradigm utilized in Experiments 9 and 10, there is also evidence for the acquisition of oculomotor action-effect associations using the same two-phase action-effect acquisition paradigm as in the experiments of the current chapter (Herwig & Horstmann, 2011). Second, numerous studies reported empirical evidence of action-effect learning in the manual domain while mostly being based on the same or slight variations of the paradigm as implemented in Experiments 9 and 10 (e.g., Camus et al., 2018; Dignath et al., 2014; Elsner & Hommel, 2001; Herwig et al., 2007). However, not every experiment utilizing the two-phase action-effect acquisition paradigm returned significant results. For example, when participants acted in a stimulus-based acquisition mode, no indication for ideomotor learning has been found for both oculomotor and manual actions (Herwig et al., 2007; Herwig & Horstmann, 2011). These findings, together with the presented results of Experiments 9 and 10, suggest that the two-phase action-effect acquisition paradigm is generally less sensitive to reveal action-effect learning than other paradigms, regardless of the action modality (i.e., whether eye movements or key presses are performed). A potential reason might be that the RT disadvantage, typically observed for the acquisition-incongruent mapping, only comes into effect when action-effect associations acquired in the preceding acquisition phase are still active. Thus, in case of quickly dissipating action-effect associations and/or the ability to quickly adapt to the context of the test phase, the two-phase action-effect acquisition paradigm might be insensitive to reveal action-effect learning that has taken place during acquisition, especially for eye move-

ments (see end of *Section 5.3.3* for an elaborate discussion). The task-irrelevant prime paradigm (as utilized in Experiments 2–8 of this work), in contrast, integrates acquisition and test phase, whereby action-effect associations are re-instantiated on a trial-by-trial basis. It seems conceivable that less memory capacity is necessary to activate action-effect associations due to this sustained reactivation, eventually increasing the sensitivity to reveal action-effect associations.

Pfister (2019) provides an alternative explanation for negative findings in studies using ideomotor paradigms. Following his rationale, one cannot rule out that participants have formed action-effect associations existing of body-related (e.g., the “feeling” of moving the eye or the tactile refference while pressing a key) rather than environmental-related action effects (e.g., visual or auditory effects). Pfister argues that ideomotor theory does not dictate which type of available action effect is used to represent an action, which leads to the conclusion that negative findings in ideomotor studies do not contradict the ideomotor idea but rather help to identify situations in which participants rely on body- (proximal) versus environmental-related (distal) action effects to represent their actions. With respect to the results of Experiments 9 and 10, it seems reasonable that participants represented the to-be-executed action, namely to perform a saccade/key press, in terms of (body- or environmental-related) action effects other than the visual on-screen action effects. In line with this, the development of paradigms capable of disentangling whether representations in terms of body- or environmental-related action effects have been used for action control could be the subject of future studies to validate these considerations. To conclude, the results of Chapter 5 provide interesting and useful insight into the nature of action-effect learning, even though the results did not show “classic” ideomotor effects.

## 6 ACTION-EFFECT COMPATIBILITY IN A GAZE LEADING TASK

While the findings presented so far have provided considerable evidence that both social and non-social consequences of oculomotor actions are integrated into saccade control, the design as implemented in Experiments 2–8 is subject to some restrictions with respect to the implications that can be drawn from these results. First, the experiments presented so far relied on an adapted version of the action-effect acquisition paradigm, which is capable to show that actions and their subsequent visual effects become associated and that these associations can be activated by the perception of the corresponding effects. Strong evidence for the ideomotor assumption of the necessity of anticipating action effects prior to action generation is still pending for gaze responses.<sup>16</sup> Second, the task as implemented in these experiments required participants to perform instructed saccades that triggered gaze responses in on-screen faces. It is fair to say that humans use their gaze quite often to achieve their goals, for example, exerting an influence on the gaze behavior of others, as it is the case when guiding the attention of an interaction partner toward an interesting object (gaze cueing; see Frischen & Tipper, 2004). However, the effect of the externally-triggered eye movements as performed by the participants in Experiments 2–8 of this work was task-irrelevant and did not serve the overall goal of establishing eye contact with the face or directing the attention of the face toward a specific location. That is, one can interpret the experimental situation in the sense that participants were in control of the gaze response of another's gaze,

---

<sup>16</sup> Note that the spatial saccade analyses of Experiment 1 and 9 have already revealed that anticipation processes occur prior to action planning, but for a non-social context only.

---

since the respective gaze response on the screen predictably occurred as a consequence of the participants' own eye movements, but not in the sense of deliberately directing another's gaze to a certain point of interest. Therefore, one should be aware that different levels of oculomotor control might be involved in these scenarios (Huestegge, Herbort, Gosch, Kunde, & Pieczykolan, 2019), even though other research from the Huestegge lab indicated that mechanisms underlying incidental versus externally triggered saccades can be quite similar (Huestegge & Koch, 2013; Huestegge, Pieczykolan, & Koch, 2014). Consequently, there is a need for the development of paradigms that account for a generalization of the present findings about oculomotor control mechanisms to control mechanisms for eye movements specifically serving the goal of directing another person's attention. The following two experiments represent a first idea of how such experiments might look like. Before I go into the methodical details, I will provide a short overview of results reported in the literature that are relevant to the following experiments, that is, evidence regarding gaze following and joint attention.

## 6.1 GAZE FOLLOWING, JOINT ATTENTION, AND GAZE LEADING

Humans seem to be highly sensitive to gaze, and thus eye movements are pivotal for human communication (e.g., Emery, 2000; Kano & Call, 2014). Eyes occupy two main functions in communication: They serve for *receiving* information from as well as *sharing* information with others in various situations. This dual function of social gaze (Gobel et al., 2015) differentiates eye movements from other human communication tools, like ears or lips, such that these other tools exclusively serve for either receiving or communicating social information. Ears cannot form sounds (i.e., cannot speak), but they are the means to hear what is said by another person, while the mouth cannot listen, but people use their mouth

---

to speak. The important role of the eyes becomes even more obvious when considering that vision can take over the function of hearing in deaf people in the form of lip-reading, and the function of communication in mute people via goal-oriented glances. Consider the following situation: You observe the gaze of another person, which is not directed at you but toward another point in space. What usually happens is that you reflexively orient your attention to the direction of the other person's gaze. This phenomenon is called *gaze following* (Emery, 2000; see Frischen et al., 2007, for a review). While you derive information from the other's eyes about her/his locus of attention, the other person utilizes gaze to communicate her/his locus of attention. Imagine that you and the other person are not only orienting toward the same direction, but now also attending to the same object as a result of your attention shift in response to the other person's gaze. This phenomenon is called *joint attention* (Emery, 2000). Thus, the initiating act in a joint attention situation is an eye movement of the *gaze leader* – also called the initiator – toward an object, causing the *gaze follower* to orient attention toward the same object.

Joint attention is a naturally occurring, reflexive, and predominantly beneficial behavior in social interactions (Frischen et al., 2007). Besides the universal information-sharing nature of joint attention – for example, when the gaze leader draws the attention of the follower toward any source of threat or toward a potential interaction partner – the ability to follow joint attention cues has been shown to be important for social development particularly for children (e.g., Farroni et al., 2002; Moore, 2008). For example, a strong ability to engage in joint attention in infancy has been demonstrated to be beneficial for the development of early language skills (Morales, Mundy, & Rojas, 1998) and theory of mind skills (Charman et al., 2000). Further, joint attention allows one to draw conclusions about the goal and to predict future actions of the initiator (Baron-Cohen, 1995; Mundy, 1995; Mundy & Newell, 2007; Tomasello, 1995). For example, if two students (Student A and Student B) are studying at the university library and Student A looks at the large clock hanging on the wall

---

in front of them, this might signal that it is time for lunch break. However, only if Student B follows the eyes of Student A, Student B will be able to derive the intention of student A from this behavior – otherwise Student B will miss the information provided by Student A and probably continues studying.

So far, research on joint attention mainly focused on the investigation of processes involved in joint attention from the *follower's* point of view, using the so-called gaze cueing paradigm (Driver et al., 1999; Friesen & Kingstone, 1998; for a review see Frischen et al., 2007), which represents a modified version of the well-known Posner spatial cueing paradigm (Posner, 1980). In a prototypical study applying the Posner spatial cueing paradigm, participants are instructed to fixate the center of the screen and to respond to the onset of a visual target, which may appear to the left or right of fixation. The target is, for example, a single letter (e.g., *T* or *L*), and participants are instructed to discriminate as quickly and accurately as possible on each trial whether the target letter was *T* or *L*. Crucially, target onset is preceded by the brief presentation of a spatial cue (e.g., an arrow pointing toward the left or right) which can be valid or not with respect to the location of the subsequent target. Such spatial cueing paradigms usually report that RTs in the letter discrimination task are faster and of higher response accuracy in the valid condition (i.e., in which targets appeared at the pre-cued location) as compared to the invalid condition (i.e., in which targets appeared at the un-cued location) even in the absence of any eye movements. These findings indicate that the spatial cue triggered an attentional shift toward the cued location, which then facilitated target discrimination in the valid (vs. invalid) condition.

In gaze cueing paradigms, the arrow cue is replaced by a centrally presented face which shifts its gaze toward a specific spatial location (e.g., toward the left or right of the center). Similar to the original Posner spatial cueing paradigm, participants showed better target detection performance in terms of RTs and accuracy rates if the gaze cue was valid as compared to if the gaze cue was invalid (Driver et al., 1999; Friesen & Kingstone, 1998). It



is important to note that these studies (Driver et al., 1999; Friesen & Kingstone, 1998) did not monitor subjects' eye movements during the task. That is, it is not clear whether *overt* (i.e., orientation toward the target via observable responses, e.g., eye movements) or *covert* (i.e., orientation toward the target in the absence of observable responses) attention shifts are responsible for the observed automaticity of orienting in response to the seen gaze. However, subsequent studies, which controlled for eye movements, were able to demonstrate that eye movements only rarely occurred in gaze cueing tasks (Friesen, Ristic, & Kingstone, 2004) suggesting that covert attentional mechanisms, as studied in the traditional Posner spatial cueing paradigm, were responsible for producing the observed effects (Friesen & Kingstone, 2003).

Generally, the results of studies using the gaze-cueing paradigm suggest that another person's gaze shift triggers a corresponding rapid and reflexive shift of attention in the observer, as previously demonstrated for traditional attention cues, such as arrows. While strong gaze cueing effects have been predominantly demonstrated with cue-target SOAs of 100–300 ms (Friesen et al., 2004; Friesen & Kingstone, 1998; Tipples, 2008), it has also been shown that the orienting response to gaze cues is rather reflexive and rapid in nature by demonstrating gaze cueing effects with cue-target SOAs as short as 14 ms (Hietanen & Leppänen, 2003). Recent approaches to increase the social nature of the experimental manipulations within the gaze-cueing paradigm focused, for example, on the assessment of affective ratings (e.g., desirability of the target object; Bayliss et al., 2006) or on the impact of factors related to the social status of the gaze leader, such as familiarity (Deaner, Shepherd, & Platt, 2007) or social dominance (Jones et al., 2010).

Humans do not only follow the other's gaze, but also use their own gaze as an effective tool to cause gaze-related effects in others (see *Section 2.3.2*). In such an interaction situation, it is of special relevance for the gaze leader to monitor the outcome of her/his gaze

behavior in the other person (Schilbach et al., 2013). Thus, it seems important to also investigate cognitive processes underlying joint attention from the perspective of the gaze leader, an issue which only recently came into focus of research (see Mundy & Newell, 2007, for a review). For example, work by Bayliss et al. (2013) has targeted the affective and behavioral consequences of causing another person to follow one's own gaze using a gaze-contingent paradigm. In their study, participants chose their preferred object from a pair of objects presented to the left and right side of a computer screen. Additionally, a face was presented at the center of the screen. After an auditory stimulus, participants communicated their choice by moving their eyes to the preferred object. The authors manipulated whether the face responded to the participant's eye movement either with a congruent or with an incongruent gaze, whereas the face identity consistently predicted whether the face would respond in a congruent or incongruent manner. Their results indicated that having one's gaze followed or not systematically affected gaze and choice behavior. For example, over the course of the experiment participants spent more time looking at the faces that never engaged in joint attention, suggesting that these faces could have been perceived as norm violators that require deeper processing in social interactions (cf. Chiappe, Brown, & Dow, 2004). In addition, the reorienting reaction toward the central face after the gaze response of that very face was influenced by congruency, demonstrating faster reorienting for the congruent as compared to the incongruent condition. Interestingly, the affective face ratings were also impacted by the congruency manipulation, such that faces consistently engaging in joint attention were preferred over the non-engaging faces, even though participants did not consciously notice the congruency manipulation. Also, participants' preference decisions were more consistent in the congruent (vs. incongruent) condition.

In summary, the gaze-contingent gaze leading paradigm as developed by Bayliss et al. (2013) was a seminal first approach to investigate behavioral consequences of engaging in joint attention. In a similar vein, recent work has shown that faces which follow people's eye

---

gaze capture attention as reflected in the efficient detection of targets presented at the follower's face (Edwards et al., 2015). The findings by Dalmaso, Edwards, and Bayliss (2016) have suggested that experiencing joint attention in terms of both gaze following and gaze leading has the power to shape subsequent social interaction processes. Moreover, in line with the finding that participants explicitly report feelings of control (Pfeiffer et al., 2013) and of realistic experience (Bayliss et al., 2013) in such gaze leading scenarios, it has been shown that initiating joint attention in a gaze leading episode generates an implicit sense of agency (for a definition of sense of agency, see *Section 4.4.3*) over the other person's congruent gaze shift (Stephenson et al., 2018).

The studies reviewed above (Bayliss et al., 2013; Dalmaso et al., 2016; Edwards et al., 2015; Stephenson et al., 2018) are major contributions to the understanding of how humans experience, comprehend, and evaluate the social consequences of their leading eye movements in a joint attention episode. To the best of my knowledge, however, the cognitive mechanisms that bring about the initiating action of the gaze leader in a joint attention episode are still unknown. The following two experiments were designed to gain insight into this particular issue and are therefore located at the crossroad of research on intentional joint attention<sup>17</sup> (Emery, 2000) and sociomotor action control (Kunde et al., 2018) in a gaze leading context. It is particularly the focus on the *initiator* of a joint attention scenario which bridges the gap between these two psychological research streams. In principle, a typical gaze leading episode can be interpreted as a specific situation in sociomotor action control, where an action – an eye movement toward an interesting object – evokes an intended perceptual action effect in another person, that is, the gaze response of the other person toward the object of interest. Thus, both theoretical approaches have in common that the gaze leader

---

<sup>17</sup> While most joint attention states are initiated on purpose by the gaze leader to encourage another person to follow her/his gaze to the intended object, it is important to note that joint attention situations also emerge by accident in everyday life, probably caused by rather intrinsic, low-level visual processing and attentional mechanisms (Frischen et al., 2007).

---

pursues the intention to lead the follower's gaze toward an interesting object. An important theoretical framework for the functional underpinnings of goal-oriented actions in social environments is the sociomotor approach to action control, which proposes that the cognitive control of such actions involves the mental anticipation of the bodily responses they evoke in other people. Experiments 3, 5, 7, and 8 of this work already provide reliable evidence for anticipation-based eye-movement control in a gaze-interaction scenario. However, the experimental setup used for these studies came with some methodological limitations, for example with respect to the assessment of anticipation processes and the ecological validity of the task, entailing a limited generalizability of the results (see *Section 6.4* for a more elaborate discussion). By implementing a joint attention episode within the action-effect compatibility paradigm, the present study intended to expand the results obtained so far in multiple ways: by creating an experimental scenario that allowed for a replication of the findings of Experiments 3 and 5 within another paradigm, and by directly assessing whether an eye movement's consequences are recollected during saccade planning, while embedding the experimental task within a more realistic gaze leading scenario to increase ecological validity. In contrast to the previous experiments in this work, the following approach represents a stronger, more natural form of controlling another's gaze with our own gaze. Participants are instructed to *guide* the gaze of another person to a certain object by moving their eyes (instead of being instructed to saccade toward a face target which responded with an arbitrary change in gaze direction). Doing so was the first approach to transfer the action-effect compatibility paradigm to the oculomotor domain – an action domain which is supposed to play a decisive role in sociomotor action control (Kunde et al., 2018).

Thus, the experiments within the present chapter were designed to address the following research question: Are eye movements toward an object triggering a gaze response in a virtual interaction partner, which would or would not correspond with the participant's gaze direction, subject to action-effect compatibility, as investigated by Kunde (2001) for the

---

manual domain. To answer this question, I conducted two experiments based on the action-effect compatibility paradigm which differed among other things in terms of whether participants acted in a forced-choice (Experiment 11) or free-choice mode (Experiment 12). Observing action-effect compatibility effects would directly corroborate the sociomotor claim that anticipated action effects in terms of a partner's gaze response play a substantial role in action selection, more specifically, in initiating joint attention.

## 6.2 EXPERIMENT 11

The procedure of Experiment 11 was a modified replication of the manual action-effect compatibility paradigm of Kunde (2001), but transferred to a social context in the oculomotor domain. Participants were instructed to respond to the gender of a centrally presented face with left- or rightward saccades toward one of two peripherally presented objects. As a result of the participant's saccade, the gaze of the on-screen face shifted, in different blocks, either toward the looked-at object or toward the opposite, currently not looked-at object. Thus, the effect of leading the gaze of the on-screen face either corresponded (compatible action-effect mapping) or did not correspond (incompatible action-effect mapping) with the saccade response. Consequently, spatial location in terms of the locus of visual attention represented the overlapping dimension for the action (i.e., saccade of the participant) and the action effect (i.e., the gaze response of the on-screen face). High effect saliency as well as an active intention to produce specific action effects have been shown to promote the anticipation of action effects (Ansorge, 2002; Janczyk, Yamaguchi, Proctor, & Pfister, 2015; Pfister & Kunde, 2013). In order to ensure that participants actually attended to the gaze response of the face and to increase effect saliency, I additionally implemented some catch-trials within each experimental block, such that participants were required to detect deviant action effects, for example, when the face responded with a gaze

shift to the left in case of a rightward gaze shift of the participant, given a compatible action-effect mapping.

I expected to observe compatibility-related effects in saccadic RTs, that is, increased saccade latencies (and/or error rates) for an incompatible (vs. compatible) action-effect mapping. In line with the reasoning in Kunde (2001), and additionally assuming that both manual and oculomotor action control are based on similar cognitive processes, such an increase would be due to the fact that the anticipated effect of a required response would additionally activate the motor pattern associated with a spatially corresponding (but currently not required) response. Observing such an effect would imply that anticipated action effects, in terms of gaze responses of an interaction partner, are important in action selection in the oculomotor domain. The effect of action-effect compatibility is usually particular evident in slow responses, since there is more time for the endogenously activated effect codes to interfere with the codes of the actually required action (Keller & Koch, 2006; Kunde, 2001; Kunde et al., 2011; Pfister, Janczyk et al., 2014; Wirth et al., 2016). Therefore, I additionally expected that compatibility effects grow in size with increasing RTs.

## 6.2.1 METHODS

### 6.2.1.1 Participants

Sample size was calculated based on the observed effect sizes in the study by Kunde (2001) with respect to the crucial compatibility effect. A power analysis using G\*Power (Faul et al., 2007) revealed a sample size of 14 participants as sufficient to observe a compatibility effect of this size with considerable power (power = .95,  $\alpha$  = .05). Nevertheless, the sample size was increased to  $N = 32$  in order to ensure a high-powered study. The 32 participants (mean age = 21.6 years,  $SD = 3.3$  years, age range: 18 – 35 years; 2 male) received either course credits or payment for compensation. All participants reported normal or corrected-

to-normal vision, and gave their informed consent before the study started. They were naïve regarding the purpose of the study.

### 6.2.1.2 Stimuli

The face photographs from Experiment 2 served as stimuli. I selected one male and one female model (see *Figure 8*, panel A for the female model), which were both available in the following variants: eyes closed, direct gaze, and averted gaze (to the left/right). The size of the face ellipses amounted to  $4.4^\circ \times 5.7^\circ$  of visual angle (maximum width  $\times$  maximum height), and faces were presented at the center of the screen. The picture of a red apple served as saccade target ( $2.2^\circ \times 2.2^\circ$  of visual angle), and appeared to the left and right ( $8.8^\circ$ ) of the centrally presented face.

### 6.2.1.3 Procedure

The experimental procedure (see *Figure 19*) was based on the design developed by Kunde (2001). Each trial started with a fixation interval of 700 ms duration displaying a white central fixation cross ( $0.4^\circ \times 0.4^\circ$  of visual angle) on a black screen. Subsequently, one of the faces (male/female with eyes closed) was displayed centrally, together with two apples in the periphery serving as saccade targets. After a variable SOA of 750 – 1250 ms (750/917/1083/1250 ms), the central face opened its eyes which served as a go-signal for the participants to shift their gaze as quickly as possible to the left or right toward one of the apple targets. The specific direction should be determined based on the gender of the face stimulus: For half of the participants, the female face required a leftward and the male face a rightward saccade, while this gender-direction mapping was reversed for the other half of participants. A correct saccade response triggered a gaze shift in the centrally presented face to the left or right side, in line with the current action-effect mapping. The gaze shift occurred with a delay of 600 ms after the saccade had landed within the interest area around the targeted apple stimulus. The gaze shift was displayed for 1,000 ms, and the next trial started

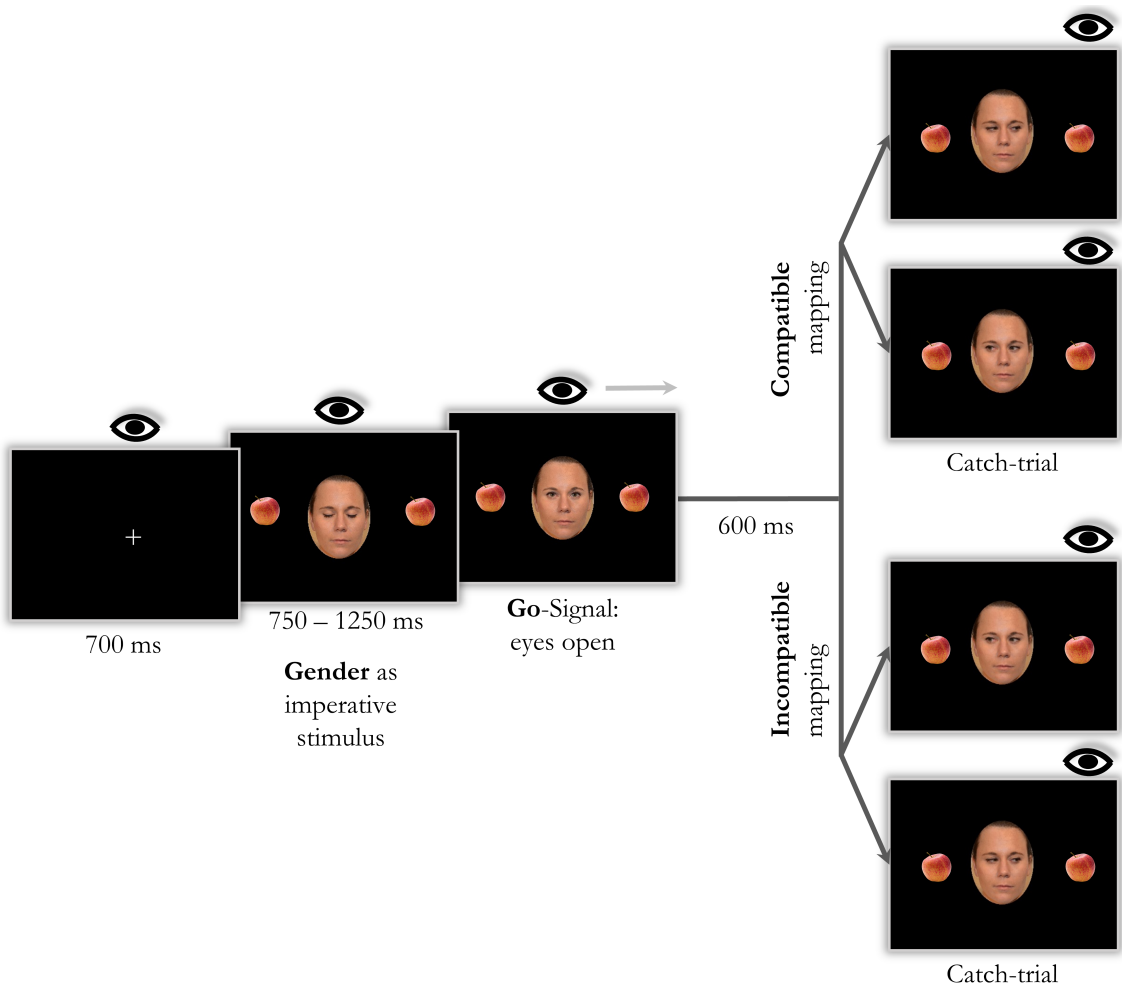
---

after an inter-trial interval (ITI) of 1,000 ms. Participants were instructed to maintain fixation on the object after their gaze leading saccade until the next trial started. In case of an erroneous saccade, the word “*Fehler!*” (German for “*Mistake!*”) was presented for 1,000 ms as soon as the erroneous saccade was detected. If no saccade to one of the two peripheral target areas was detected within 1,500 ms after the face had opened its eyes, an omission feedback message appeared for 1,000 ms displaying “*Bitte schneller reagieren und genau auf das Objekt schauen!*” (German for “*Please respond faster and fixate the object precisely!*”). In case of an error or omission trial, the next trial started immediately after the respective visual feedback had disappeared.

Participants were instructed to respond as quickly and accurately as possible as soon as the central face “had opened its eyes”. Participants were told that each correct saccade would trigger a gaze shift in the central face according to a certain rule which would be announced at the beginning of each block. In blocks with compatible action-effect mapping, it was the rule that the central face would respond with a gaze shift to the same object the participant was looking at. In blocks involving an incompatible action-effect mapping, it was the rule that the central face would respond with a gaze shift to the object opposite to where the participant was looking. I implemented catch-trials such that rule violations occurred in two randomly drawn trials within each block (one catch-trial per each model). A rule violation was defined as a gaze response of the central face into the opposite direction as defined by the active rule. Participants were instructed to respond to these rule violations by pressing the space bar. If the catch-trial was correctly identified within 2,000 ms after the onset of the incorrect gaze shift, the message “*Gut gemacht!*” (German for “*Well done!*”) was presented for 1,000 ms, and the next trial started immediately after the feedback message had disappeared. If no key press occurred within 2,000 ms, a feedback message displaying “*Hier lag eine Regelverletzung vor und Sie hätten die Leertaste drücken müssen!*“ (German for “*There was a rule*



*violation here, and this would have required to press spacebar!*") occurred. Participants were instructed to press the space bar to move on to the next trial, which started after an ITI of 1,000 ms.



*Figure 19.* Schematic representation of a trial in Experiment 11: After the presentation of a fixation screen, a central face stimulus (closed eyes) and two peripheral apples appeared simultaneously. With onset of the go-signal (eyes open), participants were instructed to shift their gaze (see symbolic eye above each screen) to the left or right apple, with the direction being determined by the gender of the face stimulus (e.g., female face = saccade to the right apple). Correct saccades triggered the onset of an action effect in terms of a gaze shift of the central face toward (in blocks with compatible action-effect mapping) or away from the targeted apple (in blocks with incompatible action-effect mapping). The action effect occurred 600 ms after a saccade had landed within the interest area around the targeted apple. In case of catch-trials, the gaze response of the central face was into the opposite direction as defined by the active action-effect mapping. Participants were instructed to respond to catch-trials by pressing the space bar.

In total, participants worked through 240 trials with a compatible and 240 trials with an incompatible action-effect mapping. Trials of each action-effect mapping condition were presented in eight consecutive blocks of 30 trials each. A calibration of the eye tracker was performed prior to each block. The order of the action-effect mapping was counterbalanced

---

across participants such that half of the participants started with the compatible action-effect mapping and received the incompatible mapping afterwards, while the order of mapping was reversed for the other half of participants. An equal number of trials with male/female stimuli was presented within each block, appearing in random order. The experiment duration amounted to around 50 minutes.

#### 6.2.1.4 Design and analysis

The present experiment involved the within-subjects factor *action-effect mapping* (compatible vs. incompatible) and the group factor *order of mapping* (compatible first vs. incompatible first). Saccade latency was defined as the interval between the onset of the imperative stimulus (i.e., when the central face opened its eyes) and the initiation of a saccade with an amplitude of at least one third of the distance between the fixation cross and the center of the object (apple). I calculated mean saccade latencies for correct trials only. Apart from the two-way ANOVAs for saccade latencies and error rates, I conducted the following analyses as suggested by Kunde (2001). First, to gain additional insight into the temporal mechanisms underlying the compatibility effect, I conducted a distribution analysis on the RT data. To do so, the RT distributions for the compatible and incompatible action-effect mapping were computed separately for each participant. In a next step, the data of each distribution were segregated into five bins. A repeated-measures ANOVA with *bin* (1–5) and *action-effect mapping* (compatible vs. incompatible) was conducted. Second, the stability of the compatibility effect was assessed by restricting the comparison of saccade latencies for compatible versus incompatible action-effect mapping to data from only the last 32 trials within each action-effect mapping.<sup>18</sup> Trials were considered erroneous if a saccade toward the left (right) target occurred when the gender-direction mapping would have required to saccade toward the right

---

<sup>18</sup> The decision to include not only the data of the last block within each action-effect mapping (corresponding to 30 trials), but to analyze data from the last 32 trials within each action-effect mapping in the stability analysis was done to align with the procedure of Kunde (2001), where the stability analysis was based on the last 32 trials within each action-effect mapping.

(left) saccade target. Apart from saccade latencies and error rates, I also analyzed key press latency and error rates for the catch-trials. Key press latency referred to the temporal interval between the onset of the gaze shift of the central face and the key press of the participant indicating that the rule violation was recognized. The catch-trial error rate referred to the number of trials where participants missed to indicate that a rule violation had occurred relative to the total number of catch-trials per each participant.<sup>19</sup>

## 6.2.2 RESULTS

Prior to analysis, all trials involving blinks, anticipatory saccades (latency < 60 ms), upward and downward saccades (instead of left/right saccades toward the target), or the absence of any saccade meeting the response criteria, as outlined above, were removed. This procedure resulted in the exclusion of 2.1% of all trials, which corresponded to 350 out of 15,360 trials in total (32 participants with 480 trials each).

### 6.2.2.1 Saccade latencies

Mean saccade latencies for the compatible versus incompatible action-effect mapping amounted to 345 ms ( $SE = 18$  ms) versus 324 ms ( $SE = 15$  ms) when participants started with the compatible action-effect mapping, and to 310 ms ( $SE = 18$  ms) versus 325 ms ( $SE = 15$  ms) when participants started with the incompatible action-effect mapping (see *Figure 20*). The two-way ANOVA with action-effect mapping and order of mapping revealed no significant main effects, both  $F$ s < 1, but a significant interaction,  $F(1, 30) = 12.88$ ,  $p = .001$ ,  $\eta^2_p = .30$ . While participants responded faster with compatible (vs. incompatible)

---

<sup>19</sup> Please note that the total number of catch-trials varied between participants. The reason for this is that I excluded trials from analysis where no saccade toward one of the objects had been detected within a certain time interval and no gaze response of the virtual face occurred. If this happened in catch-trials, the total number of catch-trials was reduced for the respective participant.

---

action-effect mappings when they encountered the incompatible action-effect mapping condition first,  $p = .043$ , the pattern was reversed when they encountered the compatible action-effect mapping first,  $p = .006$ .

### 6.2.2.2 Distribution and stability analysis

The distribution analysis of the RT data indicated no significant effect other than the (self-evident) effect of bin,  $F(4, 124) = 290.04$ ,  $\varepsilon = .36$ ,  $p < .001$ ,  $\eta^2_p = .90$ , all other  $F$ s  $< 1$  (see *Figure 20*). Including the factor order of mapping within the distribution analysis revealed a significant three-way interaction,  $F(4, 120) = 49.01$ ,  $\varepsilon = .36$ ,  $p < .001$ ,  $\eta^2_p = .62$ , demonstrating that the compatibility effect increased with increasing RTs in opposite directions for the two order conditions. Pairwise comparisons revealed a significant effect of compatibility in terms of faster RTs for compatible (vs. incompatible) action-effect mappings only in the last bin ( $p < .001$ ) in the *incompatible first* condition, while the effect of compatibility indicated slower RTs for compatible (vs. incompatible) action-effect mappings from the third bin on ( $ps < .05$ ) in the *compatible first* condition. None of the other statistical comparisons involving the factor order was significant, all  $F$ s  $< 1$ . To assess the stability of the effect, I conducted a repeated-measures ANOVA with action-effect mapping and order of mapping, which did not reveal any significant results, all  $F$ s  $< 1$ .

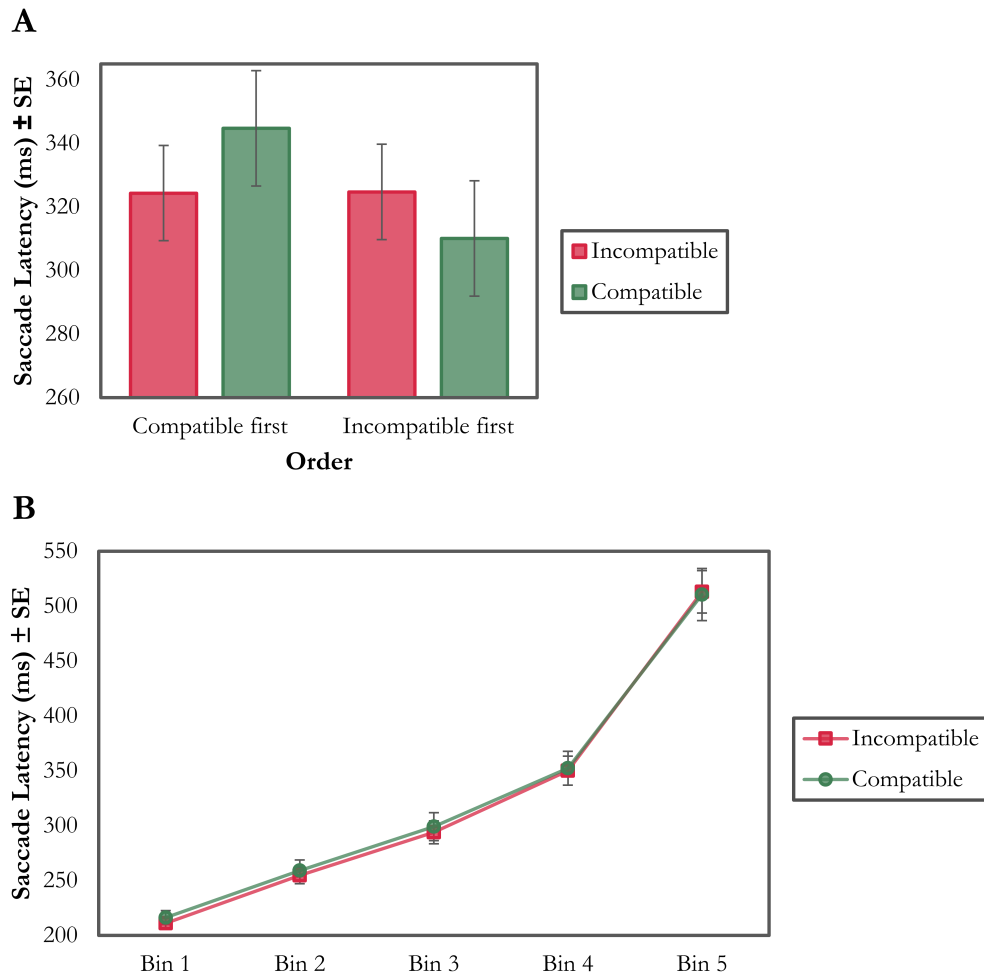


Figure 20. Mean saccade latencies (in ms) in Experiment 11 (A) as a function of action-effect mapping (compatible vs. incompatible) and order of mapping (compatible first vs. incompatible first), and (B) as a function of action-effect mapping (compatible vs. incompatible) and bin (1–5), averaged across order of mapping conditions. Error bars depict standard errors of the mean ( $SE$  in ms).

### 6.3.2.3 Oculomotor error rates

The mean error rate for the compatible versus incompatible mappings amounted to 1.2% ( $SE = 0.3\%$ ) versus 1.3% ( $SE = 0.4\%$ ) for participants who experienced the compatible action-effect mapping first, and to 0.9% ( $SE = 0.3\%$ ) versus 1.6% ( $SE = 0.4\%$ ) for participants who experienced the incompatible action-effect mapping first. The repeated-measures ANOVA with action-effect mapping and order of mapping did not show any significant effects, all  $F$ s  $< 1$  (except for the effect of action-effect mapping,  $F(1, 30) = 1.04$ ,  $p = .316$ ,  $\eta^2_p = .03$ ).

#### 6.2.2.4 Catch-trial analysis

A paired-samples  $t$ -test revealed no significant difference in catch-trial key press latencies between compatible ( $M = 681$  ms,  $SE = 21$  ms) and incompatible ( $M = 700$  ms,  $SE = 26$  ms) action-effect mapping conditions,  $t < 1$ . Similarly, the difference in error rates for compatible ( $M = 8.8\%$ ,  $SE = 1.9\%$ ) versus incompatible action-effect mapping conditions ( $M = 13.1\%$ ,  $SE = 2.4\%$ ) was not significant,  $t(31) = 1.56$ ,  $p = .129$ ,  $d = .28$ .

### 6.2.3 DISCUSSION

The purpose of this experiment was to investigate whether the spatial correspondence between eye movements toward an object and the resulting gaze reaction of an on-screen face toward or away from the attended object has an impact on the ease of generating these saccades. Observing such an influence in terms of compatibility effects would provide direct evidence for the core ideomotor assumption that anticipated action effects are pivotal for action generation – a conclusion which can not directly be drawn from studies relying on the action-effect acquisition paradigm (see *Section 2.1.2* for details about the different paradigms).

The results showed that participants' responses toward an object were faster when the other's gaze followed their gaze toward the object compared to when the other's gaze was directed in opposite direction, but only for the subgroup of participants who experienced the incompatible mapping first. This finding replicates – at least partially – the basic action-effect compatibility effect reported by Kunde (2001) and suggests that humans do anticipate another's gaze responses to their own gaze when selecting these actions. These anticipations are, according to ideomotor hypothesis, assumed to facilitate performance in the group with compatible mappings due to a dimensional overlap between the features of the actual movement (e.g., looking toward the left object) and the features activated by the anticipation (e.g., the other person looking toward the left object) and to impair performance for the group

with incompatible mappings. These results are a first indication of a generalization of the action-effect compatibility effect to the oculomotor domain as investigated in a gaze leading episode.

Even though these results at first sight appear promising by suggesting that anticipated action effects play an important role in action selection in a more applied gaze-leading scenario, the results should be taken with caution. First, evidence of anticipation was sequence-dependent: While the subgroup experiencing the incompatible action-effect mapping first showed action-effect compatibility effects as predicted, this effect was reversed in the compatible-first group. A similar result pattern was evident for verbal action-effect compatibility effects (cf. Koch & Kunde, 2002, Experiment 2; see Experiment 12 of the present work for further discussion). Moreover, the results of the stability analysis suggest that the observed data pattern might be explained by a “surprise” hypothesis, as called by Kunde. According to this idea, effects of action-effect compatibility originate from violations of pre-experimentally established action-effect associations, which remove cognitive resources from processing the task, rather than from anticipated action effects as suggested by ideomotor theory. Following this rationale, compatibility effects should primarily emerge from the initial trials with incompatible action-effect mapping and vanish over the course of the experiment. This was indeed the case in the present experiment: In contrast to previous findings (Kunde, 2001; Kunde et al., 2004), the compatibility manipulation had no influence on saccade latencies when only the last few trials of each action-effect mapping were analyzed. It might be that the already rather weak influence of action-effect compatibility observed in Experiment 11 fails to be reflected in RTs after several trial repetitions, possibly due to reduced predictability and fast adaptability of eye movements to situational contexts. One could assume that the mental representation of the anticipated action effect, that is, the left or right gaze shift of the on-screen face, primes the locus of attention (“left” or “right”) and thereby introduces a compatible or incompatible relationship between the to-be-executed action and the action

---

effect, but this action-effect compatibility is not strong enough to affect saccade performance after several trial repetitions.

The design of the present experiment comes with a major constraint, which renders it likely that a potential impact of action-effect compatibility did not fully emerge. More specifically, participants in the present study saw a centrally presented female or male face with initially closed eyes that opened after a variable SOA. While the opening of the eyes served as the start signal to initiate a response saccade, gender was the feature defining the direction of that saccade (to the left or right object). Consequently, participants could have prepared their action prior to the onset of the go-signal while the face, and therefore the direction-defining gender, was already visible. This implies that RTs measured for saccades initiated after the go-signal might not reflect processes of action selection (as intended by the action-effect compatibility paradigm), but of action initiation only.

Similar problems also arise in the context of free-choice tasks, where participants could also theoretically select and prepare one of the action alternatives before onset of the go-signal, which would also reduce the impact of action-effect anticipation. This objection can be partially dispelled with studies showing that compatibility effects also emerged for pre-cued, and thus fully prepared actions (Kunde et al., 2004), albeit such effects appear to be considerably smaller compared to conditions preventing response selection to take place before the onset of the go-signal (for a related finding see also Elsner & Hommel, 2001, Experiments 2 and 3). To overcome this pitfall and to further validate the results of Experiment 11 (in one of the order conditions), I conducted Experiment 12, which differed from Experiment 11 in several major points while being based on the same action-effect compatibility paradigm. Another possibility would be to just cancel the “closed-eyes screen” and re-run Experiment 11. However, I decided against that option but rather conducted an experiment which was improved not only with respect to the “closed-eyes screen”, but also with respect to a more natural gaze leading scenario, which will be described in the following.



---

## 6.3 EXPERIMENT 12

Experiment 12 was also based on the action-effect compatibility procedure (Kunde, 2001), but here it was further adapted in order to resemble more a natural setting of a gaze leading episode, where the leading person usually pursues the internal goal of directing the gaze of another person toward a certain object of interest, thereby establishing joint attention (Frischen et al., 2007). Thus, I embedded the task in a cover story taking place at a casino scenery, which will be described in detail in *Section 6.4.1*. In brief, participants played the role of a trick cheater with the intention to train a student cheater on how to cheat at card games. They were told that they should imagine to be in a training situation, and that their task was to convey to their student the information about the next intended action in a card game with the help of their eyes such that no other player at the gaming table will recognize it. More specifically, participants saw the face of their student presented centrally on the screen, surrounded by two card decks to the left and right and the picture of a lamp presented at the top of the screen. They were instructed to respond to the onset of the lamp with either a left- or a rightward saccade toward one of the card decks depending on their choice. As a consequence of their saccade, the student responded (in different blocks) with a gaze shift which either corresponded or did not correspond with the teacher's response. That is, the student's gaze shift was directed either toward the card deck chosen by the teacher (compatible action-effect mapping) or toward the opposite card deck (incompatible action-effect mapping). The compatible and incompatible action-effect mapping were motivated based on two different strategies to prevent the casino from detecting the trickery (i.e., sometimes the teacher looks to the side opposite to the one side the student should attend to in order to make sure that the casino does not detect a consistent gaze communication strategy). Additionally, participants were instructed to monitor their student's performance and to indicate

when their student violated the current rule (see below for more details). Participants received extra monetary compensation as an incentive to attend to the student's performance.

The chosen design of Experiment 12 comes with several advantages. First, the action-effect compatibility design is implemented in an even more natural setting where participants use their gaze in order to lead the gaze of their student. Second, and apart from the use of a cover story, the key difference to Experiment 11 was that participants acted in a free-choice mode. While such free-choice procedures have been shown to slightly attenuate the action-effect compatibility influence in the manual domain (Kunde, 2001, Experiment 3), they represent a method to rule out that observed action-effect compatibility effects are the result of an influence of acquired stimulus-effect instead of action-effect associations (see Elsner & Hommel, 2001, and Kunde, 2001, for a more elaborate discussion). Like in Experiment 11, I implemented some trials where the student's gaze response deviated from the currently valid mapping (i.e., catch-trials). Participants were instructed to select their actions spontaneously, but to execute each response about equally often and in a non-systematic order as soon as the go-signal occurred. In order to prevent premature response decisions, that is, that action selection occurred prior to the onset of the go-signal, I implemented no-go trials where participants were asked to withhold any action. Again, I expected to observe an influence of action-effect compatibility in terms of faster saccade latencies for compatible versus incompatible action-effect mappings, an effect that should be more pronounced with increasing saccade latencies.

## 6.3.1 METHODS

### 6.3.1.1 Participants

In total, 62 participants were recruited. One participant was excluded from the analysis because of missing data in at least one cell of the experimental design, and another one was excluded due to unusually high RTs ( $> 3 SDs$ ). Data of the remaining 60 participants

were analyzed (mean age = 24.5 years,  $SD = 4.1$  years, age range: 18 – 36 years; 11 male). They received either course credits or payment for compensation. All participants reported normal or corrected-to-normal vision. Before the experiment started, they gave informed consent. They were naïve with respect to the purpose of the study.

### 6.3.1.2 Stimuli

In Experiment 12, only the face of the female model (available in the following variants: direct gaze and averted gaze to the left/right) from Experiment 11 served as a stimulus. The face stimulus was presented at the center of the screen. The picture of two card decks (maximum width  $\times$  maximum height of each picture:  $5.3^\circ \times 3.5^\circ$  of visual angle) served as saccade targets. These saccade targets appeared at a distance of  $8.8^\circ$  of visual angle to the left and right of the central face. The picture of a lamp ( $3.5^\circ \times 3.1^\circ$  of visual angle) that was presented above the central face served as an imperative stimulus, and was available in the following variants: light off and yellow/red light on (see *Figure 21*).

### 6.3.1.3 Procedure

Experiment 12 made use of a cover story in order to increase external and ecological validity of the results. I told participants that cheating must be learned and that they would play the role of a teacher who trains a student on how to cheat in a card game. To do so, participants were told that they had to guide the attention of their student toward a specific card deck from which the next card should be drawn with the help of their eye movements. To indicate that the student had registered the teacher's intention, he would shift his gaze toward the card deck from which he would draw the next card. In order to avoid that the casino would become aware of the cheating, participants were told that they agree on two different strategies with their student: Looking at a card deck would mean to draw a card from the looked-at card deck for the first strategy, while it would mean to draw a card from the opposite card deck (i.e., the not-looked-at card deck) for the second strategy. They were

---

told that the information which strategy will be applied is presented before each block. They further received the information that the student would make mistakes from time to time, and that their task was to indicate erroneous behavior via a key press (space bar).

Participants were instructed that the casino employees would particularly closely look at the gambling table via monitors from time to time. If this was the case, the lamp would turn red and neither an eye movement nor any other action must be executed by the participant. Finally, participants were informed that they could earn extra compensation in terms of more money for good performance if they detect the student's mistakes and correctly withhold any gaze response in case of the red lamp. This incentive system was implemented in order to increase the participants' motivation to pay attention to the student's gaze shift (i.e., the action effect).

More precisely, each trial started with the presentation of a white fixation cross ( $0.4^\circ \times 0.4^\circ$  of visual angle) on a black screen (see *Figure 21*). After that, the central face (gaze directed at the participant), the two card decks to the left and right in the periphery, as well as the lamp (light off) appeared on the display. After a variable SOA of 750 – 1250 ms (750/917/1083/1250 ms), the lamp turned yellow (red) indicating a go (no-go) trial. In go trials, participants were instructed to shift their gaze as quickly as possible toward one of the two (left/right) card decks as soon as the yellow light appeared. Participants were instructed to choose spontaneously to look to the left or right deck in response to the go signal, but to look at each card deck about equally often and in random order. Before the experiment started, it was emphasized during the instruction that they should focus more on spontaneous decisions rather than on a perfectly even distribution of responses. Only if the distribution of choice frequencies of left- and rightward saccades deviated significantly from the instructed balanced distribution, that is, if more than 21 saccades to the left/right occurred within each block of 33 trials, on-screen feedback regarding the imbalance was provided to the participant at the end of the respective block. In this case, the on-screen feedback also

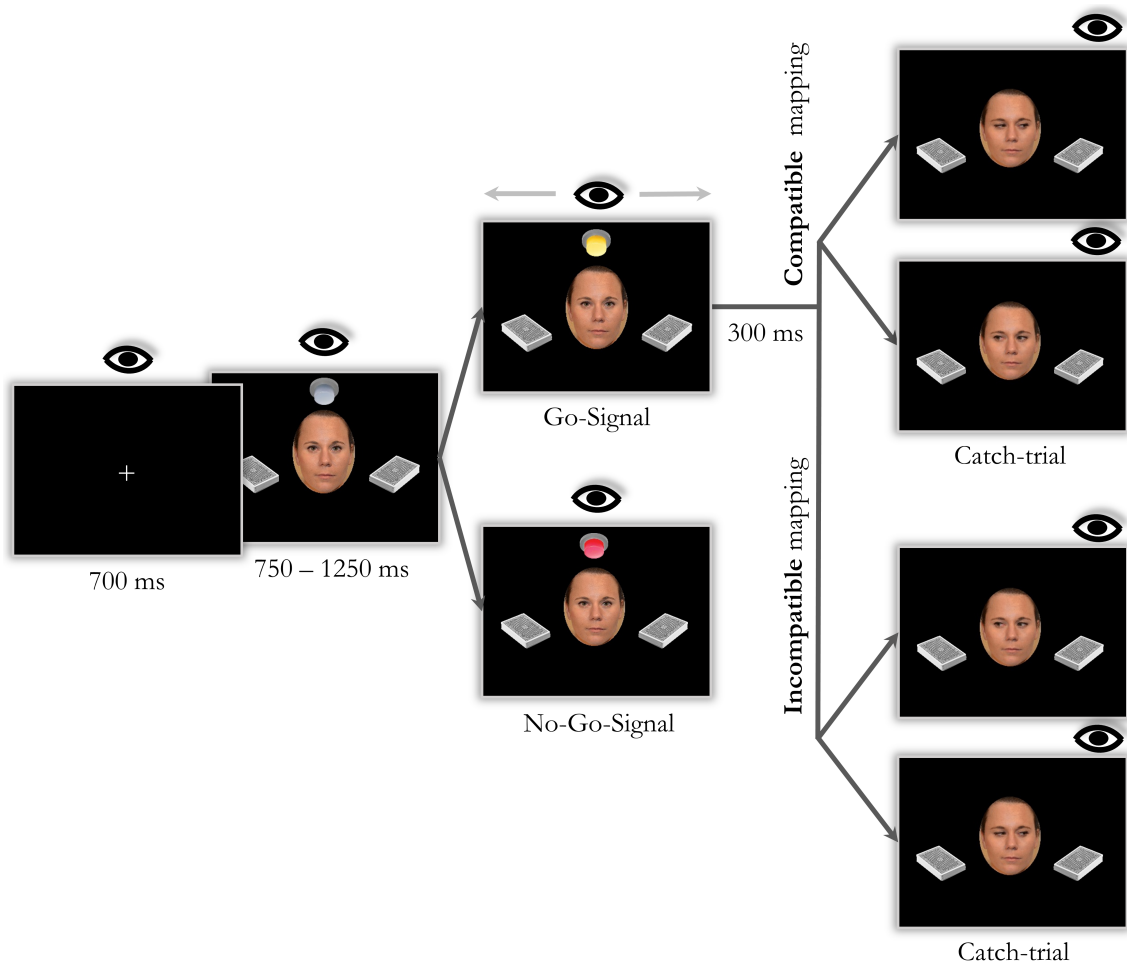
included a reminder to choose each direction about equally often, and participants could proceed to the next block via pressing the space bar. In no-go trials, participants were instructed to withhold any (oculomotor and manual) response. One randomly drawn trial out of eleven trials was a no-go trial.

Each saccade to one of the card decks triggered the central face to shift its gaze either to the left or to the right card deck, according to the current action-effect mapping. In case of a compatible action-effect mapping, the central face shifted its gaze to the looked-at card deck, whereas in the incompatible action-effect mapping, the central face shifted its gaze to the opposite card deck (i.e., the not-looked-at card deck). The delay between the landing of a target saccade within an interest area around one of the card decks and the onset of the gaze shift of the central face was 300 ms. The gaze shift of the central face was visible for 1,000 ms, and participants were instructed to fixate on the card deck until the next trial started (after an ITI of 1,000 ms). If 1,500 ms had elapsed without detecting any saccade toward one of the peripheral card decks, an omission feedback message was presented (1,000 ms) displaying "*Bitte schneller reagieren und genau auf das Objekt schauen!*" (German for "*Please respond faster and fixate the object precisely!*"). The next trial started immediately after the omission feedback had disappeared. Analogous to the catch-trials implemented in Experiment 11, the student's response was erroneous in one randomly drawn trial out of eleven trials. Participants were told to respond to these erroneous responses by pressing the space bar. If the space bar was pressed within 2,000 ms after the onset of the gaze shift, the message "*Gut gemacht!*" (German for "*Well done!*") was presented for 1,000 ms, and the next trial started immediately after the feedback display had disappeared. If no key press occurred within 2,000 ms, a feedback message displaying "*Hier lag ein Fehler vor und Sie hätten die Leertaste drücken müssen!*" (German for "*There was a mistake here, and this would have required to press the space bar!*") occurred. Additionally, the feedback message "*Hier lag kein Fehler vor und Sie hätten die Leertaste nicht drücken dürfen!*" (German for "*There was no mistake here, and you should not have pressed space bar!*")

---

was displayed for 1,000 ms if the space bar was pressed in case of regular trials (i.e., in trials in which the student did not commit any mistake). In the latter two cases, participants were instructed to press the space bar to move on, and the next trial started after an ITI of 1,000 ms. A post-survey at the end of the experiment asked participants to indicate the subjectively perceived difficulty of the task (i.e., to monitor the student's performance) with the help of a nine-point scale ranging from 1 (very easy) to 9 (very difficult).

Participants encountered 264 trials of each action-effect mapping, resulting in a total number of 528 trials. Each mapping condition consisted of eight blocks with 33 trials each, and three catch-trials and three no-go trials were presented within each block, while one catch-trial and no-go trial each were randomly presented within every sequence of eleven trials. A calibration of the eye tracker was performed prior to each block. The experiment duration was around 60 minutes. The order of action-effect mappings was counterbalanced across participants: Half of the participants started with four blocks of the compatible action-effect mapping (compatible mapping = CM), followed by eight blocks of the incompatible action-effect mapping (incompatible mapping = IM), and finished with the remaining four blocks of the compatible action-effect mapping afterwards, resulting in the sequence CM – IM – IM – CM. The other half of participants received this sequence in reversed order (IM – CM – CM – IM) (see Rieger, 2007, for using similar sequences). The use of these sequences comes with the advantage of accounting (at least partly) for a frequently changing context in social situations (as compared to the design of Kunde, 2001, where the action-effect mapping was reversed in the middle of the experiment only). Moreover, this sequential design also allows for a separate analysis of the data from the first and second half of the experiment (involving 132 trials in each compatibility condition).



*Figure 21.* Illustration of the trial structure in Experiment 12: After the presentation of a fixation screen, a centrally presented face, two peripheral card decks, and a lamp (top of the screen) appeared simultaneously. With onset of the go-signal (lamp turning yellow), participants shifted their gaze to the left or right card deck (free-choice), triggering the onset of an action effect in terms of a gaze shift of the central face toward (in blocks with compatible action-effect mapping) or away from the targeted card deck (in blocks with incompatible action-effect mapping). The action effect occurred 300 ms after a saccade had landed within the interest area around the targeted card deck. In case of catch-trials, the gaze response of the central face was into the opposite direction as defined by the active action-effect mapping. Participants were instructed to respond to catch-trials by pressing the space bar. In case of a no-go signal (lamp turning red), participants were instructed to withhold any (oculomotor and manual) response.

#### 6.3.1.4 Design and analysis

The present experiment involved the within-subjects factor *action-effect mapping* (compatible vs. incompatible) and the group factor *order of mapping* with the values compatible first

---

(CM – IM – IM – CM) and incompatible first (IM – CM – CM – IM).<sup>20</sup> Saccade latency was defined as the interval between the onset of the yellow lamp and the initiation of a saccade with an amplitude of at least one third of the distance between the fixation cross and the center of the card deck in the present experiment. I analyzed saccade latencies in the same way as in Experiment 11 (see *Section 6.2.1.4*). Additionally, the design of Experiment 12 allowed for an analysis of the effects of *action-effect mapping* and *order of mapping* separately for the first versus second half of the experiment. Catch-trials, key press latencies and error rates were analyzed in the exact same way as described in Experiment 11.

### 6.3.2 RESULTS

Prior to analysis, all trials involving blinks, anticipatory saccades (latency < 60 ms), or the absence of any saccade meeting the response criteria as outlined above were removed. Further, I excluded all trials where no saccade landed within one of the two the pre-defined areas around the card decks. No-go trials were also excluded prior to analysis. This cleansing procedure resulted in the exclusion of 8.0% of all trials, which corresponded to 2530 of the total number of 31,680 trials (60 participants with 528 trials each, including no-go and catch-trials). The average proportion of left- versus rightward saccades of all trials amounted to 49.1% versus 50.9%, respectively. Thus, the distribution of choice frequencies of response saccades to the left and right card deck did not significantly deviate from the instructed balanced distribution, as indicated by a non-significant statistical comparison,  $t(61) = 1.59$ ,  $p =$

---

<sup>20</sup> Please note that the factor order of mapping has a slightly different meaning in Experiment 12 as compared to the previously presented experiment: In Experiment 11, the values compatible/incompatible first refer to the fact that all trials corresponding to the compatible/incompatible action-effect mapping were experienced *before* the action-effect mapping was reversed in the second half of the experiment. In Experiment 12, however, the respective values refer to the action-effect mapping which has been experienced during the very first block of the experiment. When only the second half of Experiment 12 is analyzed, being in the *compatible first* order condition actually means that the first block of the second part of the experiment was a block with incompatible action-effect mapping (and vice versa).



.118,  $d = .21$ . No participant had to be excluded due to an unbalanced choice frequency distribution (a ratio equal to or exceeding 2:1).

### 6.3.2.1 Saccade latencies

Saccade latencies were submitted to a two-way repeated-measures ANOVA with action-effect mapping and order of mapping as within-subjects factors. Mean saccade latencies for the compatible versus incompatible action-effect mapping amounted to 394 ms (SE = 12 ms) versus 397 ms (SE = 13 ms) when participants experienced the compatible action-effect mapping first, and to 360 ms (SE = 12 ms) versus 370 ms (SE = 13 ms) when participants started with the incompatible action-effect mapping (see *Figure 22*). The effect of action-effect mapping was significant,  $F(1, 58) = 4.66, p = .035, \eta^2_p = .07$ . Neither the interaction of action-effect mapping and order of mapping,  $F(1, 58) = 1.45, p = .233, \eta^2_p = .02$ , nor the main effect of order were significant,  $F(1, 58) = 3.26, p = .076, \eta^2_p = .05$ .

### 6.3.2.2 Distribution and stability analysis

Distribution analysis of the RT data using a repeated-measures ANOVA with bin and action-effect mapping as factors indicated an increase of the action-effect compatibility effect with increasing bin,  $F(4, 236) = 3.82, \varepsilon = .38, p = .036, \eta^2_p = .06$ , for the interaction of action-effect mapping and bin (see *Figure 22*). Pairwise  $t$ -tests revealed a significant influence of action-effect mapping in the last bin,  $p = .019$ , and a close to significant effect in the fourth bin,  $p = .066$ , while no such influence was apparent in the remaining bins, all  $p$ s > .125. For the sake of completeness, I will also report the two main effects. Both, the effect of action-effect mapping,  $F(1, 59) = 4.64, p = .035, \eta^2_p = .07$ , and the effect of bin,  $F(4, 236) = 624.92, \varepsilon = .29, p < .001, \eta^2_p = .91$ , were significant.

In order to ensure comparability between Experiments 11 and 12, the stability analysis of Experiment 12 was based on data from the last 32 trials within each action-effect mapping from the *first* half of the experiment. By doing so, only trials from the first encounter

with each action-effect mapping were considered in the stability analysis. The repeated-measures ANOVA with action-effect mapping and order of mapping yielded no significant effect of action-effect mapping,  $F < 1$ . The effect of order of mapping was significant,  $F(1, 58) = 4.18, p = .046, \eta^2_p = .07$ , which was further specified by the significant interaction of action-effect mapping and order of mapping,  $F(1, 58) = 7.16, p = .010, \eta^2_p = .11$ . Saccade latencies were higher when encountering the compatible (vs. incompatible) mapping first, but only for the compatible action-effect mapping,  $p = .006$ , while no such differences were evident for the incompatible action-effect mapping,  $p = .338$ .

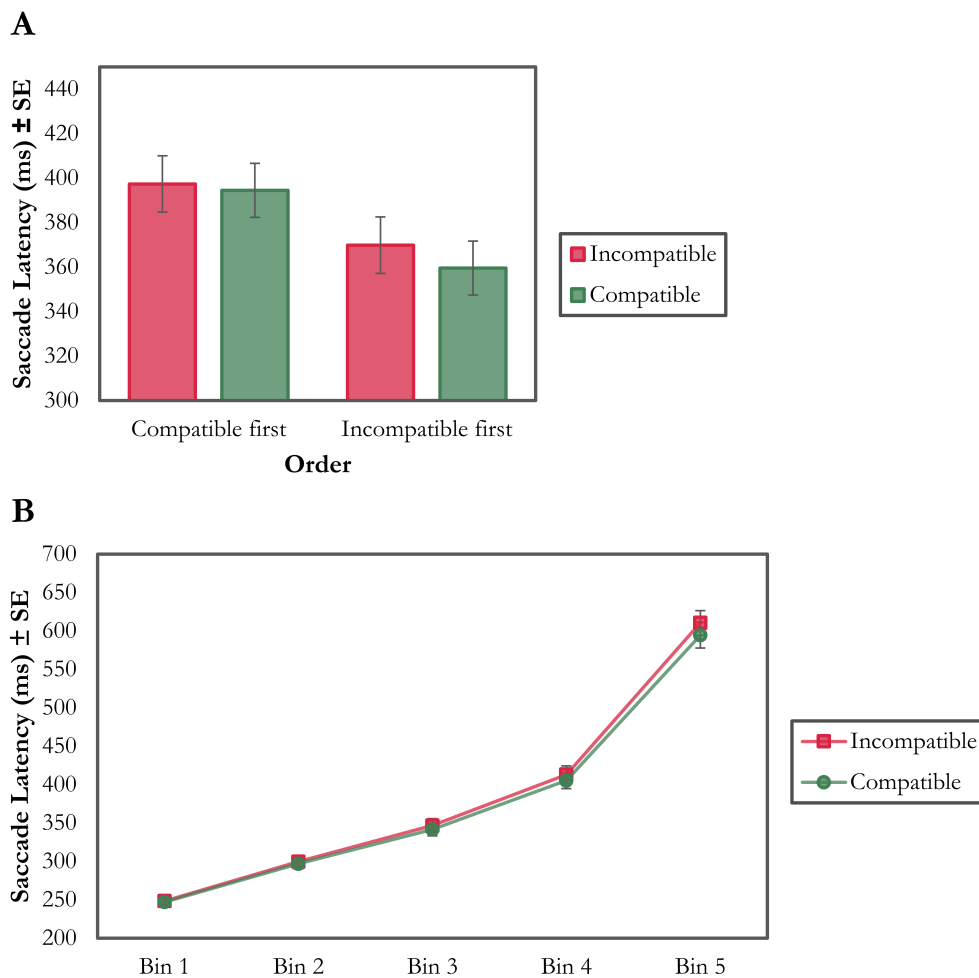


Figure 22. Mean saccade latencies (in ms) in Experiment 12 (A) as a function of action-effect mapping (compatible vs. incompatible) and order of mapping (compatible first vs. incompatible first), and (B) as a function of action-effect mapping (compatible vs. incompatible) and bin (1–5), averaged across order of mapping conditions. Error bars depict standard errors of the mean ( $SE$  in ms).

### 6.3.2.3 Effect of experiment half

A three-way repeated-measures ANOVA with the additional within-subjects factor *experiment half* (first vs. second) was conducted for saccade latencies. For the sake of readability, I only report statistical comparisons involving the factor experiment half. Participants responded faster in the first ( $M = 368$  ms,  $SE = 9$  ms) versus second ( $M = 393$  ms,  $SE = 9$  ms) half of the experiment,  $F(1, 58) = 22.30$ ,  $p < .001$ ,  $\eta^2_p = .28$ . Both the interaction of experiment half and order of mapping,  $F < 1$ , and the interaction of experiment half and action-effect mapping,  $F(1, 58) = 2.97$ ,  $p = .090$ ,  $\eta^2_p = .05$ , yielded no significant effects. However, the three-way interaction was significant,  $F(1, 58) = 22.38$ ,  $p < .001$ ,  $\eta^2_p = .08$ . Separate two-way ANOVAs conducted for each experiment half demonstrated an influence of action-effect mapping (in terms of faster RTs for compatible vs. incompatible action-effect mapping) in the first half of the experiment, but only when participants experienced the incompatible action-effect mapping first,  $F(1, 58) = 15.78$ ,  $p < .001$ ,  $\eta^2_p = .21$ , for the interaction of action-effect mapping and order of mapping. In the second half of the experiment, the interaction of action-effect mapping and order of mapping was significant, too,  $F(1, 58) = 6.59$ ,  $p = .013$ ,  $\eta^2_p = .10$ , indicating an influence of action-effect mapping only when the first block (of the entire experiment) was a block with compatible action-effect mapping. In terms of the two different orders of mapping-sequences as implemented in the present study, this actually means that a compatibility effect was present in the second half of the experiment when the compatible mapping followed the incompatible mapping.

### 6.3.2.4 Catch-trial analysis

Key press latency for catch-trials was not significantly different for the compatible ( $M = 611$  ms,  $SE = 15$  ms) versus incompatible action-effect mapping condition ( $M = 602$  ms,  $SE = 16$  ms),  $t(59) = 1.16$ ,  $p = .253$ ,  $d = .15$ . Similarly, the error rate for catch-trials did not significantly differ between the compatible ( $M = 5.2\%$ ,  $SE = 0.7\%$ ) and incompatible action-effect mapping condition ( $M = 6.1\%$ ,  $SE = 0.8\%$ ),  $t < 1$ .

### 6.3.3 DISCUSSION

Experiment 12 aimed at a more consistent demonstration of the effect observed in Experiment 11 by further increasing the ecological validity of the experimental setting. Additionally, the experimental design of Experiment 11 did not make it possible to exclude that premature action selection may have reduced the impact of anticipated action effects (see *Section 6.2.3*). Therefore, I embedded the action-effect compatibility paradigm into a cover story, putting the participant in the position of a person who teaches another person how to cheat in the casino, and implemented a free-choice task. The incorporation of no-go trials further prevented premature action selection to occur. Hence, with Experiment 12, I tested whether the ease of generating an eye movement toward an object depended on whether one's gaze was subsequently followed or not. This was indeed the case: A significant influence of action-effect mapping was observed, indicating that eye movements toward an object triggering another person to follow with her/his gaze toward the same object are easier to initiate than eye movements toward an object causing another person to direct her/his gaze in the opposite direction. Additionally, the impact of action-effect compatibility increased with higher RTs as observed by Kunde (2001). However, this impact did not remain stable with practice as indicated by the absence of a compatibility effect in the stability analysis. In sum, Experiment 12 replicated and extended the main results of Experiment 11, and the results are more consistent in the present experiment.

At first sight, action-effect compatibility did not seem to be affected by mapping order in Experiment 12, which replicates the results of Kunde (2001), but contrasts the results of Experiment 11, where the compatibility influence was shown to be sequence-dependent. However, the sequential design I implemented in the present Experiment also allowed for an analysis of the data separately for the first and second half of the experiment. This analysis revealed a sequential modulation of the compatibility influence. Compatibility

effects in terms of an RT advantage for a compatible (vs. incompatible) action-effect mapping were always evident when a block with compatible mapping followed a block with incompatible mapping, regardless of whether this was the first encounter with both mappings or not. This finding is of special interest when viewed in conjunction with the opposing compatibility influence in Experiment 11 depending on which mapping was experienced first. While the results of Experiment 11 might solely be explained as being practice-related (i.e., faster responses for the mapping that was experienced in the second half of the experiment), this explanation does not hold for the results of Experiment 12, where effects of action-effect compatibility emerged despite participants being highly familiar with both mappings. On the contrary, the results indicate that experiencing an incompatible action-effect mapping is rather effortful and not to the same extent prone to exercise effects as the compatible mapping. This interpretation is supported by an observation from the stability analysis of the *first half* of the experiment where RTs for the incompatible mapping condition were shown to be comparable for the last trials of the incompatible mapping regardless of the order of mapping, whereas RTs for the compatible mapping condition were significantly slower when the compatible condition was experienced first (as compared to when the compatible condition was experienced second). A suggestion to eliminate (or reduce) possible sequence effects is to introduce some practice trials of each action-effect compatibility mapping before the actual experiment starts in order to reduce possible effects of surprise (when encountering an action-effect mapping for the first time) or habituation.

Note that the study of Edwards et al. (2015) revealed that people rapidly orient their attention toward an individual who has followed their gaze and thus has established joint attention. In the current experiment, however, I wanted to prevent participants from re-directing their gaze toward the gaze follower because that would introduce another compatibility dimension, namely between the gaze response of the follower and the direction of the participant's re-orienting saccade, which would be exactly in opposition to the compatibility

with respect to the looked-at object. Consequently, participants received the instruction to maintain their gaze at the gazed-at object until the next trial started. Additional analyses showed that some participants disregarded the instruction to maintain fixation. Excluding all “instruction refusers”, that is, all participants who oriented their gaze back to fixation in more than one quarter of the trials (corresponding to 13 participants), left the overall result pattern unchanged.<sup>21</sup> However, a comparison of the effect size of the compatibility effect for the whole sample ( $\eta^2_p = .07$ ) and the subset of instruction-followers ( $\eta^2_p = .14$ ) indicated a stronger compatibility effect when excluding the “instruction refusers” from analysis. This finding implies that future studies on action-effect compatibility in a gaze leading episode should probably control for uninstructed eye movements after the target saccade has been carried out or focus specifically on investigating the effects of such reorienting saccades on action-effect compatibility effects.

## 6.4 CONCLUSION OF CHAPTER 6

This chapter comprises the dedicated study of gaze control in a social context using the action-effect compatibility paradigm. Therefore, I adapted the paradigm in order to apply it to a gaze leading episode. To the best of my knowledge, gaze control has not yet been studied within this paradigm, although it has considerable advantages. Most importantly, observing compatibility effects would serve as convincing evidence for the core ideomotor claim that action effects become anticipated when generating goal-oriented actions. Further, the nature of the action-effect compatibility paradigm allows for the study of goal-oriented

---

<sup>21</sup> The repeated-measures ANOVA including only the subset of participants who maintained fixation, that is, instruction-followers, revealed a significant main effect of action-effect mapping,  $F(1, 45) = 7.45, p = .009, \eta^2_p = .14$ . Saccade latencies amounted to 371 ms ( $SE = 10$  ms) versus 380 ms ( $SE = 10$  ms) for compatible versus incompatible action-effect mappings. Both the effect of order,  $F(1, 45) = 3.07, p = .087, \eta^2_p = .06$ , and the interaction,  $F(1, 45) = 1.37, p = .248, \eta^2_p = .03$ , were not significant.

---

eye movements within a scenario of higher ecological validity, in which participants deliberately direct another's gaze to a certain point of interest. In both experiments, I observed effects of action-effect compatibility (although only inconsistently in Experiment 11; for possible reasons see *Section 6.2.3*). That is, the action effects in terms of the gaze responses of another person affected saccade latencies although these effects were presented exclusively after the saccade had been initiated, which is in line with the core ideomotor assumption of an endogenous activation of action effects prior to action generation.

Distribution analyses usually demonstrate that the effects of action-effect compatibility are stronger with longer RTs (Kunde, 2001). In line with these prior findings, I observed especially strong compatibility effects for slow responses in Experiment 12, and in the subgroup of participants who encountered the incompatible condition first in Experiment 11. This is in line with the idea of Kunde (2001), who suggested that there is more time for the time-consuming endogenous activation of the effect codes in the case of longer RTs, eventually causing stronger interference with the codes of the actually required action. However, a closer look at the nature of the corresponding interaction effect observed in Experiments 11 and 12 points toward some differences when compared to the results of Kunde (2001) and other studies. While compatibility effects were shown to be evident throughout all bins or starting from the second bin and to increase with increasing RTs in the study of Kunde (2001), effects of compatibility here were restricted to the last bin(s) in Experiments 11 and 12 of this work. However, the absolute RTs still reveal a tendency toward faster RTs for compatible as compared to incompatible conditions even in earlier bins, although the statistical comparisons were not significant. It is important to note that these differences between the results of Kunde (2001) and the results of the current experiments do not undermine the underlying rationale in general, but give rise to the presumption that compatibility effects are less pronounced for oculomotor (vs. manual) actions. More critical for the interpretation presented above is the observation that I observed a reversed compatibility

influence that predominantly emerged from slow responses in the subgroup of Experiment 11 who experienced the compatible mapping first (see *Section 6.2.3*).

The compatibility effects observed in Experiments 11 and 12 are characterized by rather small effect sizes as compared to those usually obtained from studies in the manual domain, for example with body-related action effects (Wirth et al., 2016) or with complex actions (Janczyk et al., 2015). In addition to the small effect size, the stability analysis of both presented experiments revealed that the compatibility effect dissipated over the course of the experiment, a finding that contrasts with the rather robust effect observed by Kunde (2001). In comparison with studies from the social domain, however, the effect size is similar to what has been observed in previous studies, for example when studying effects of imitation versus counter-imitation in an action-effect compatibility task (Pfister et al., 2017) or effect-based control of facial expressions (Kunde et al., 2011).

This observation allows three conclusions to be drawn which are not mutually exclusive and which could be the subject of future studies. First, effects of action-effect compatibility might generally be weaker when investigated in the social domain due to the more ambiguous and less predictable nature of social environments. It seems conceivable that an anticipated gaze response has a reduced impact on performance, since the acting person knows from everyday learning experiences that the other person could respond to her/his gaze with various options: for example, by following her/his gaze toward the relevant object (as intended by the actor), by starting to talk to the actor, or by ignoring the actor's gaze at all. Second, action selection in the present tasks was rather simple and detached from a real-life interaction context, which is usually characterized by the fact that one can select from more than two alternative action goals. Thus, effects of anticipated gaze responses might be more pronounced when a choice between several action goals is possible. Finally, the action effects in the form of gaze responses were probably a bit artificial when compared to gaze responses observed in real life (despite greater ecological validity compared to Experiments



1–8). For the sake of simplicity and experimental control, the action effects were implemented in the form of visual changes in photographs of faces rather than in the form of gaze responses of real interaction partners that would contain all naturally occurring movement cues. This deliberation is related to the discussion whether social behavior can be captured by the presentation of pictures of faces at all, instead of interacting with real people, which is covered in the General discussion of this work (see *Section 7.3.1*).

An interesting observation from research on effect monitoring is that incompatible action effects are generally harder to monitor than compatible action effects as reflected in higher dual-task costs (Wirth, Janczyk, & Kunde, 2018). Applying this logic to the experiments presented in this chapter, one could assume that participants should be able to detect catch-trials more accurately while encountering the compatible rather than the incompatible action-effect mapping. However, the presented data do not support this conclusion, as indicated by the non-significant results of the *t*-test comparing catch-trial error rates between the compatibility conditions, even though the numerical values of both experiments (and especially of Experiment 11) show a trend toward the assumed effect. The absence of significant differences may be due to the fact that the relevant action effect was presented at the visual periphery, which is known to increase processing time and impair accuracy as compared to foveal presentation (Bayle, Schoendorff, Hénaff, & Krolak-Salmon, 2011; Carrasco, Evert, Chang, & Katz, 1995; Jüttner & Rentschler, 2000; Wolfe, O'Neill, & Bennett, 1998; but see Carrasco, McElree, Denisova, & Giordano, 2003, for a counter-example, where processing speed increased with increasing eccentricity). Further, the visual changes in the form of a movement of the eyeball from straight gaze to averted gaze (to the left or right), which served as action effects, were of rather low visual saliency. It may be the case that participants in the present study experienced the task to monitor the peripherally occurring action effect rather difficult, which might have masked differences between the compatibility conditions. This suggestion is corroborated by anecdotal evidence from the post-experiment survey, in which

some participants reported that they had a hard time to detect the gaze changes in general. However, despite these subjective difficulties to detect the gaze shift, the average error rates were rather low. To sum up, future studies should consider these deliberations, possibly by developing an experimental setup where the action effects are of greater visual saliency and presented at reduced eccentricity.

---

# 7 GENERAL DISCUSSION

## 7.1 SUMMARY OF THE RESULTS

Gaze is an important and powerful tool for communication with others. Therefore, this work aimed at the investigation of cognitive mechanisms underlying intentional eye-movement control in a gaze-interaction scenario. While there was already evidence for effect-based eye-movement control from contexts other than gaze interaction (Herwig & Horstmann, 2011; Huestegge & Kreutzfeldt, 2012; Pfeuffer et al., 2016), a goal throughout this work was to replicate and expand these findings to a gaze-interaction scenario. Another purpose was to study whether such kind of eye-movement control is qualitatively different from eye-movement control in the non-social domain. It was proposed that both gaze control in a social and a non-social context is subject to effect-based control processes, but given that social (vs. non-social) gaze control might come with some peculiarities inherent to the less predictable and more ambiguous character of the social environment (cf. Kunde et al., 2018), I explored whether the underlying cognitive mechanisms differ for gaze control in a social (vs. non-social) context.

The findings presented in this work can be summarized by three key messages. First, I observed that participants could acquire bi-directional action-effect associations between their saccades and the subsequently perceived gaze response of another person in a gaze-interaction context, which in turn affected oculomotor control. These results show for the first time that eye movements in a gaze-interaction scenario are represented in terms of their

---

gaze response in others. Second, on a more general level, the presented series of experiments confirms that the congruency effect in saccade latencies – first described by Huestegge and Kreutzfeldt (2012) for the non-social domain using the same one-phase action-effect acquisition paradigm as implemented in Experiments 2–8 – represents a relevant and consistent finding, since this effect has been replicated several times with different stimuli within this work. However, the observed congruency effect is subject to the limitation that the action effect either must come with sufficient visual saliency (see Experiments 1, 3, 7, and 8) or must be integrated into a meaningful social context – when rather unobtrusive (see Experiments 5, 7, 8, 11, and 12) – in order to affect saccade control. Finally, gaze control in a social environment is not special. More precisely, the results of the presented series of experiments overall speak against the assumption of fundamentally separate mechanisms underlying ideomotor gaze control in a social versus non-social environment. Instead, the findings suggest that independent of whether someone intends to affect the gaze behavior of another person or to manipulate an inanimate object with the help of her/his gaze, saccade control is based on the same general psychological mechanisms. Thus, rather than opposing ideomotor and sociomotor gaze control, the results of this work suggest to consider sociomotor gaze control as an integral part of ideomotor gaze control. In the following paragraph, I will briefly summarize the individual results of the empirical chapters, before continuing with a general discussion.

The first empirical part of this dissertation (Chapter 3, Experiment 1) demonstrated that the anticipation of oculomotor action effects is reflected in spatial saccade characteristics in inanimate environments, similar to corresponding effects observed in a social context (Herwig & Horstmann, 2011). Moreover, the combined results of Experiments 2–6 (Chapter 3) demonstrated congruency effects (as an indicator of the acquisition of bi-directional associations between saccades and their effects) for both social and non-social stimuli, but only when the perceived changes are sufficiently salient to be incorporated into saccade control.

---

Together, these results suggest similar general psychological mechanisms underlying ideomotor gaze control in a social versus non-social context. In the following empirical part (Chapter 4, Experiments 7 and 8), I investigated whether oculomotor action-effect learning is subject to associative learning mechanisms known to affect action-effect learning in the manual domain. These results revealed oculomotor action-effect learning to be rather unaffected by substantially reducing effect contingency, and to be most pronounced when action and effect are separated by a temporal delay of 200 ms. In line with the observations of Chapter 3 (Experiments 1–6), oculomotor action-effect learning here again did not depend on the type of environment (social vs. non-social). The third empirical part (Chapter 5, Experiments 9 and 10) addressed the role of different action modes for the acquisition of action-effect associations for both oculomotor and manual actions. However, no congruency effects were observed in neither oculomotor nor manual RTs, therefore no conclusion about the role of different action modes during acquisition on action-effect learning was feasible. The last empirical part (Chapter 6, Experiments 11 and 12) aimed at transferring the previous findings on saccade control in a gaze-interaction scenario to an experimental design with higher ecological validity, which was especially suited to test whether planning of a saccade involves the anticipation of the targeted gaze response of the other person. Both experiments (Experiment 12 in particular) indeed showed effects of effect anticipation in a gaze-leading task, suggesting that anticipations are at play when generating saccades in order to control the gaze of another person.

---

## 7.2 TOWARD A MODEL OF SHARED COGNITIVE MECHANISMS OF OCULOMOTOR ACTION CONTROL

### 7.2.1 SPECIFICATIONS OF THE MODEL

The observed results in this work support recent theoretical claims emphasizing the role of anticipation-based action control in social interaction (Kunde et al., 2018), while no fundamental differences to anticipation-based action control in a non-social context were identified. Thus, I will present a model of shared cognitive mechanisms underlying oculomotor action control in both social and non-social environments which is based on the pioneering work on ideomotor action control (e.g., James, 1890) and the latest developments of ideomotor action control (e.g., Huestegge & Kreutzfeldt, 2012; Kunde et al., 2018). Huestegge and Kreutzfeldt (2012) were among the first to study the impact of associations between oculomotor actions and their visual effects in saccade control, thereby transferring the ideomotor idea to the oculomotor domain (together with Herwig & Horstmann, 2011). They proposed that goal-oriented eye movements are represented in terms of their intended effects in the environment. Kunde et al. (2018) established a theoretical framework for ideomotor action control in a social context (sociomotor action control), proposing that social actions (including social eye movements) are represented (just like non-social actions) in terms of their intended effects in the environment, that is, effects evoked in other people. However, a systematic comparison of ideomotor eye-movement control in a social versus non-social environment with respect to a potentially special role of gaze in social contexts (see *Section 2.3.2*) was still pending. This work provides empirical evidence for the parsimonious theoretical rationale of similar processes underlying effect-based gaze control in a social as compared to a non-social action context. This key finding of the presented work is captured in the following model of shared cognitive mechanisms of anticipation-based oculomotor action control (see *Figure 23*).

---

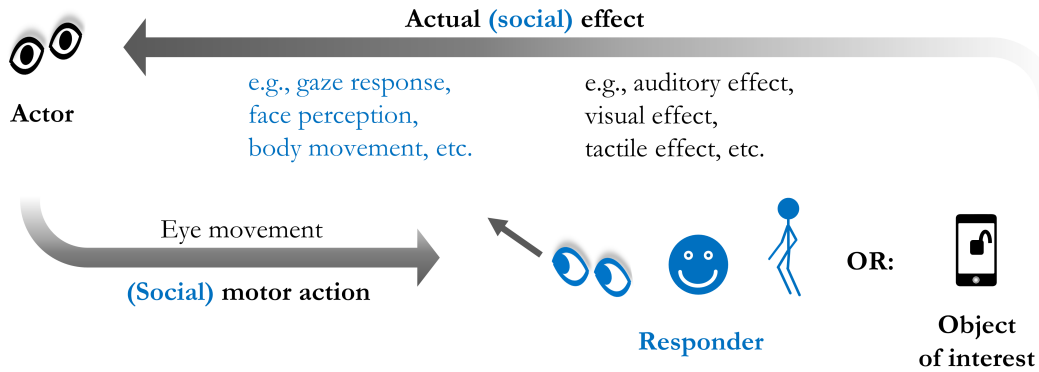
The proposed model assumes essentially the same cognitive mechanisms underlying gaze control independent of the character of the context (social vs. non-social). Therefore, the model suggests that eye movements in both social and non-social contexts become associated with the consequences that these eye movements consistently evoke. This action-effect learning provides the basis for the production of a certain effect in the environment – be it to affect another person or the inanimate environment. The model assumes that the activated action goal, that is, the intended action effect, reactivates the eye movement that was linked to the current action goal before. According to the proposed model, an agent selects to perform a specific saccade by representing the intended action effect, independent of whether the action goal is to draw the attention of a confederate to the winning deck of cards (Experiment 12, see *Section 6.3*) or to select an object on the screen via gaze.

It is known from the manual domain that the same motor action usually produces multiple consequences, for instance in terms of tactile, visual, and auditory reafferences, and that situational and intentional factors flexibly determine what kind of reafference eventually governs action production (Kunde & Weigelt, 2005; Memelink & Hommel, 2013). The experiments presented in this work exclusively studied visual action effects since the most basic effect of each saccade lies in the visual perception of the post-saccadic object. Whether other reafferences of goal-oriented eye movements, for instance eye movements with vibrotactile (see Rantala et al., 2020, for a review and design guidelines of gaze interaction with vibrotactile feedback in human-computer interaction) or auditory feedback, might be used in a similar way to retrieve an action still needs to be tested. However, I propose that non-visual reafferences of action effects should principally also be capable of retrieving eye movements associated with generating the intended action effect.

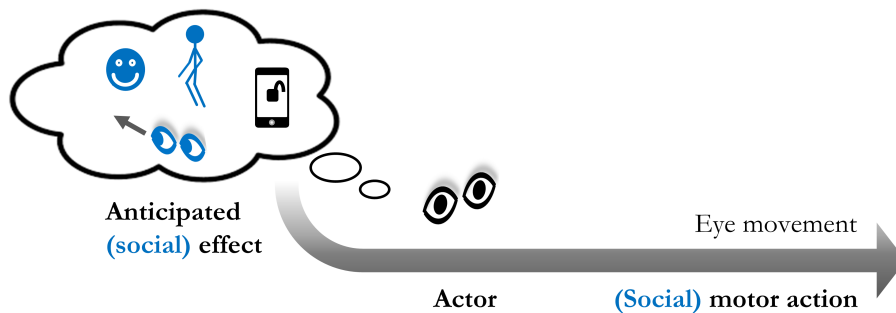
The proposed model of shared cognitive mechanisms of oculomotor action control contributes to the distinct fields of cognitive and social psychology by integrating both perspectives on human behavior into a single theoretical concept. This approach is consistent

with recent claims, for instance by Kim and Hommel (2019), that call for developing mechanistic theories of social phenomena that explain, rather than merely describe, these phenomena.

### A: Action-effect learning



### B: Action planning



*Figure 23.* Proposed model of shared cognitive mechanisms of oculomotor action control in social and non-social environments. Actors acquire knowledge about which eye movement evokes what kind of action effect in the (social or non-social) environment, thereby acquiring bi-directional action-effect associations (A). After learning, the anticipation of the intended action effect triggers the corresponding eye movement (B).

## 7.2.2 IMPLICATIONS ON EXISTING OCULOMOTOR CONTROL THEORIES

The presented results also bear implications for oculomotor control theories. Previous theories of oculomotor control mechanisms usually did not focus on anticipatory processes in eye-movement control apart from rather low-level processes associated with enabling basic perceptual stability or improving perceptual processes (e.g., Herwig & Schneider, 2014; Rolfs, Jonikaitis, Deubel, & Cavanagh, 2011). The influential model of eye-movement control, for example, proposed by Findlay and Walker (1999) postulates two parallel and



---

partially interactive processing streams, a temporal “when” and a spatial “where” stream, which contribute to the generation of saccadic eye movements. Thus, their model accounts well for the generation of low-level, stimulus-elicited saccades. Moreover, Findlay and Walker also briefly describe cognitive influences on saccade control by differentiating between different levels of cognitive control. Specifically, the model distinguishes between an automatic, an automatized, and a voluntary level of eye-movement control, while cognitive processes only play a role at the latter one. According to this model, eye movements at this voluntary control level are generated by top-down, instruction-driven weighting processes. For example, the generation of saccades in the opposite direction to a peripheral, highly-salient target (as investigated in the anti-saccade paradigm, see Hallett, 1978; Massen, 2004; Walker, Husain, Hodgson, Harrison, & Kennard, 1998) is based on the abovementioned level of voluntary eye-movement control.

The results of my work, however, corroborate the claim of Huestegge and Kreuzfeldt (2012) that not only instructions, but also the anticipated identity of the intended action effect plays an important role for saccade generation at a voluntary control level (independent of the social vs. non-social nature of the anticipated action effect), and should therefore be considered explicitly in theories of oculomotor control. Consequently, adding a “what” processing stream to the model of Findlay and Walker would provide a more complete picture of cognitive processes underlying oculomotor control. Huestegge et al. (2019), for instance, have taken a first step in that direction by proposing a model of eye-movement control which is based on the same control levels as suggested by Findlay and Walker (1999), but focusing on free-choice oculomotor behavior as a type of voluntary eye-movement control which has been scarcely considered in previous models. In their model, automatic, automatized, and voluntary control levels are assumed to interact in a bi-directional manner with control principles operating on either lower (e.g., saliency-based) or higher (e.g., free-choice) levels of

---

control, thereby determining saccade generation. Some recent studies also addressed the impact of reward on saccade control (see McCoy & Theeuwes, 2016). While reward represents a specific type of (valence) anticipation, which has the power to shape saccadic behavior, it only covers a single facet of anticipation. A general framework of anticipation-based oculomotor control is still pending. This work – and the proposed model of shared cognitive mechanisms of oculomotor action control (see *Section 7.2.1*) – together with the model of Huestegge et al. (2019) is therefore just the beginning of a more comprehensive understanding of oculomotor action control and might thus help to further encourage researchers to tackle this important principle of motor control.

## 7.3 OPEN ISSUES AND FUTURE CHALLENGES

### 7.3.1 MAINTAINING EXPERIMENTAL CONTROL VERSUS GENERALIZABILITY OF RESULTS

The designed experiments in this work aimed at achieving a balance between the ecological validity of the task and the stimuli combined with maintaining high experimental control. As a result of these efforts, one might argue that the experimental environment I refer to as establishing a “social context” is not sufficiently social in that it only involved pictures of faces that have no social relevance for the participant and thus do not really establish a social situation comparable to the presence of real human interaction partners. However, I think that the results of the present experiments using schematic face stimuli (Experiments 3 and 7–10) represent an important first step in this novel area of research, and demonstrate that under controlled conditions, evidence for the acquisition of links between one’s gaze and the gaze response of another (albeit virtual) person can be shown. This approach is similar to research in the related field of gaze cueing, where presenting face stimuli on a screen – either schematic (e.g., Friesen & Kingstone, 1998) or photographic (e.g.,

Driver et al., 1999) – has been considered a stimulating initial approach to the study of fundamental social cognitive processes (for a review, see Frischen et al., 2007). Nevertheless, a growing number of researchers argued that the investigation of social behavior, and social attention in particular, can never be complete without additionally considering the complexity and ambiguity of real world social situations (Birmingham, Ristic, & Kingstone, 2012; Hayward, Voorhies, Morris, Capozzi, & Ristic, 2017). Such efforts have already been made in the investigation of gaze cueing and joint attention effects (e.g., Gallup et al., 2012; Lachat et al., 2012; Redcay et al., 2010) and have been addressed in this work by investigating action-effect compatibility in a more realistic social interaction setting in Experiments 11 and 12. Both experiments yielded evidence for an important role of anticipated action effects in terms of gaze responses of an interaction partner in action selection in the oculomotor domain (albeit somewhat inconsistent in Experiment 11). These results – together with the above-mentioned approaches to investigate social cognition in more realistic scenarios – are encouraging to tackle this issue by developing suitable paradigms of even higher ecological validity to study gaze interaction in “real world” social behavior (e.g., Birmingham et al., 2012; Pfister et al., 2013; Schilbach et al., 2013; Zaki, Kallman, Wimmer, Ochsner, & Shohamy, 2016; Zaki & Ochsner, 2009). The gaze cueing effect, for example, has already been investigated in a face-to-face interaction situation as opposed to a computer-based manner (Lachat et al., 2012), demonstrating a gaze cueing effect of about the same robustness and size. However, these efforts to validate laboratory-based findings are especially important since there are also empirical findings, for example on attention-related mechanisms, that seem to differ when investigated with computer-based designs versus with real people (e.g., Laidlaw, Foulsham, Kuhn, & Kingstone, 2011; Rubo, Huestegge, & Gamer, 2020).

The effect of a social as opposed to a non-social context on gaze control has been investigated in the present work by using different pictures representing the action effects in each context type while keeping visual low-level features constant across context types (for

---

the stimulus material, see *Figure 8*). While the presented findings suggest a negligible role of context type for mechanisms of gaze control, the method used does not rule out the possibility that the absence of any differences in gaze control is due to the different stimulus material per se rather than the context manipulation (social vs. non-social). To substantiate the findings of a negligible role of context type, one should conduct further studies where identical stimuli serve as action effects in both contexts, while they are interpreted as belonging to a social context or not. For instance, one could rely on a gaze-interaction scenario similar to the one established in Experiment 12, while participants would be told that they are collaborating with another person such that the gaze response of the virtual face represents the behavior of a real person sitting in the room next door in the social context condition. In the non-social condition, participants would receive the instruction that the behavior of the virtual face is controlled by a computer.

### 7.3.2 FURTHER INSTANCES OF GOAL-ORIENTED GLANCES

Up to now, effect-based control of eye movements has been investigated using changes in visual perception as action effects, assuming that eye movements are planned in anticipation of the visual effects they evoke. The empirical setups used throughout this work to investigate the role of oculomotor action-effect associations for eye-movement control involved actions which resulted in the *onset* of any visual stimulation, for instance in terms of traffic lights turning on (Experiment 1) or gaze aversion/following (Experiments 2–12). However, there are also situations in which an eye movement is performed with the aim of preventing any action effect to occur. In this case, the intended action effect involves the *absence* of a certain visual stimulation or behavioral response, respectively. For instance, if a person approaches trying to get in contact with you, you might decide to look at the floor intentionally or turn your gaze to the side to prevent that person from getting in touch with you. In this case, the action effect would correspond to stopping the other person. At the

moment, it is not clear whether cognitive mechanisms underlying the control of this type of eye movements – aimed at preventing any certain visual input or behavioral response – are the same or different from the underlying mechanisms of eye movements intending to stimulate visual input or behavioral responses. Gaze avoidance behavior is especially pronounced in people with social anxiety disorder (Weeks, Howell, & Goldin, 2013). The findings that socially anxious people report an enhanced feeling of being looked at as compared to healthy controls (Gamer, Hecht, Seipp, & Hiller, 2011; Schulze, Renneberg, & Lobmaier, 2013) may contribute to the explanation of such kind of gaze avoidance behavior. When viewed from an ideomotor perspective of action control, the question of whether action control differs for patients with social anxiety as compared to healthy controls in situations where gaze is used to stop being looked at has not yet been addressed.

Besides the decision of what to do and when to act, intentional action control also involves the question of whether to execute any action at all (Brass & Haggard, 2008). This means that in certain circumstances it may be appropriate not to act in order to bring about an intended action effect. Just like the effects of acting, effects of not acting might either elicit perceptual changes in the non-social environment or trigger a desired behavior in any other person. For instance, imagine the signalman, who can decide which track the train will use because he has the possibility to set the points. In a given situation, the signalman could decide *not* to switch the points such that the train will reach the desired destination, and thereby performing a non-action. Further, a person who wants to increase her/his attraction to others might deliberately not respond to a message in order to increase the other person's efforts to attract attention. Intentional omission of actions, so called non-actions, have also been studied from the theoretical perspective of ideomotor action control (Kühn & Brass, 2010; Kühn, Elsner, Prinz, & Brass, 2009; Weller, Kunde, & Pfister, 2017), providing evidence for the general assumption of the ideomotor idea that actions and non-actions are in

---

principle represented and controlled in the same fashion via their anticipated sensory consequences. Even though corresponding results were obtained from the manual domain, where the decision to execute or to omit a key press produced a contingent sound (Kühn et al., 2009; Weller, Kunde et al., 2017) or an on-screen visual effect (Weller, Kunde et al., 2017, Experiment 1), these findings might also be transferrable to effect-based control of eye movements. Thus, people might from time to time decide to omit an eye movement in order to reach a certain goal. Corresponding effects could be located on a perceptual level, for instance when the intended action effect lies in the perception of any object that was previously covered by another object moving through the field of vision. More precisely, the decision not to follow the opening theatre curtain with one's gaze, but to focus one's gaze centrally on the stage, allows one to concentrate immediately on what is happening on stage. Furthermore, the desired effect of omitting an eye movement could also lie in triggering a broad range of social effects. For example, the decision not to avoid the gaze of another person during a dispute, which means staring at the opponent, could lead the other person to eventually give in.

However, this way of looking at non-actions entails terminological entanglements that can easily be resolved in the manual, but not in the oculomotor action domain. In the case of manual actions, the definition of a non-action is rather unambiguous in that it is defined as the intentional omission of an action (Weller, Kunde et al., 2017), for instance, choosing not to press a key. This definition is based on a dichotomous system, but cannot easily be transferred to the oculomotor domain. At first sight, one could simply state that not performing a saccade represents an oculomotor non-action. However, apart from the problem that our eyes can never really rest, deciding not to act opens up several alternative action contexts, all of which represent non-actions at their core, but are fundamentally different in meaning. For instance, not acting in oculomotor action control can relate to maintain fixation at the currently fixated object (e.g., staring at someone) or to not move the eyes to a certain

---

object of interest (e.g., avoid looking at someone/something). Closing the eyes could also represent a facet of oculomotor non-action. These considerations may partially explain why research on oculomotor non-actions is scarce to date. However, combining research on oculomotor action control with research on intentional non-actions represents a promising and fruitful approach to contribute to a comprehensive view of this particular ideomotor research area.

### 7.3.3 GAZE CONTROL AND TECHNICAL PROGRESS

Finally, it should be noted that a potential special role of eye movements in a social as opposed to a non-social environment (see *Section 2.3.1* and *Section 2.3.2*) is fading with technical progress. From an evolutionary perspective, it is reasonable to assume that the eyes are reserved for controlling human behavior (e.g. gaze cueing), but more and more areas are being added in which gaze as an alternative input modality for communication is becoming the norm, especially as the technology behind it becomes increasingly affordable and suitable for everyday use. The possibility to communicate in a gaze-based way is of special relevance for motor-impaired patients who might not be able to rely on traditional communication techniques such as mouse or keyboard use (Slobodenyuk, 2016). Previous research has demonstrated the applicability of eye movements as input devices in fields like drawing applications (Hornof & Cavender, 2005; van der Kamp & Sundstedt, 2011), gaming (Corcoran, Nanu, Petrescu, & Bigioi, 2012), typing (Akkil et al., 2016; Mott, Williams, Wobbrock, & Morris, 2017), or web browsing (Abe, Owada, Ohi, & Ohyama, 2008). Consequently, the gaze itself can take over several functions (see Majaranta, Riih , Hyrskykari, & Špakov, 2019, for a recent review), such as pointing (Asai et al., 2000), zooming (Adams, Witkowski, & Spence, 2008; Halwani, Salcudean, Lessoway, & Fels, 2017), or object selection (Tanriverdi & Jacob, 2000; Urbina & Huckauf, 2008). With respect to certain functions, the use of eye movements as input devices has even been shown to be superior compared to conventional

---

input devices, for example, in terms of faster pointing with an eye mark pointer (i.e., an eye-tracking-system-based sensor) as compared to a joystick pointer (Asai et al., 2000). However, the development of such interfaces raises questions with respect to the integration of the cognitive architecture of the user to maximize usability and to minimize interference with other ongoing cognitive processes (Slobodenyuk, 2016). Slobodenyuk (2016), for example, calls for the development of *cognitively grounded gaze-controlled interfaces*, taking into account the findings within the research areas related to sense of agency and embodied cognition for design considerations. The findings of my work on the cognitive foundations of goal-oriented eye movements might be helpful to root the ideomotor perspective on human action control in corresponding design considerations, for instance when it comes to understand how predictable consequences of oculomotor actions govern one's own action control.

Moreover, the present work on gaze interaction can also be relevant for the field of human-robot interaction when designing machines who are capable of responding to human gaze not only with vocal or motor responses (in terms of locomotion), but also with gaze responses. Socially assistive robotics, for instance, is a field of application in which findings from the area of cognitive psychology often find their way into design considerations for different health care interventions, such as stroke rehabilitation, mental health care, or training programs for patients with dementia or children with autism (see Matarić, 2017, for a recent review). Socially assistive robots are designed for social rather than physical interaction with the user. According to Matarić (2017) the most challenging part in designing socially assistive robotics is to integrate the complex facets of human social interaction including physical, social, and cognitive aspects into a consistent overall system in order to prevent user rejection. To sum up, given recent technical innovations, an understanding of the cognitive foundations of oculomotor actions in social and non-social contexts is probably of direct relevance for the development of gaze-controlled interfaces and human-robot interaction, and the present work represents an important step in that direction.



---

## 7.4 CONCLUDING REMARKS

Humans move their eyes constantly and rapidly, thereby executing up to four saccades per second (Liversedge et al., 2011). While the majority of these saccades are generated to explore the visual environment, some of these eye movements are initiated to pursue a specific goal and to change the environment in a desired way. The present work was dedicated to the study of these goal-oriented eye movements by focusing on goals related to changes in other human's gaze behavior, for instance, when directing another person's attention toward an object of interest. To explore the impact of a social (vs. non-social) context on basic mechanisms of human action control, eye movements with the goal to evoke desired effects in the non-social environment served as comparison conditions. The experiments aimed at the understanding of the cognitive processes that bring about such goal-oriented eye movements, and to highlight potential peculiarities that specifically shape gaze control in a social context. The theoretical rationale underlying the presented series of experiments was the ideomotor theory of action control and its extension to the social domain. The presented results showed for the first time that oculomotor actions that evoke gaze responses in others are represented and brought about in terms of their intended effects, and that these oculomotor actions share essential characteristics with those in the non-social domain. Thus, potential peculiarities of ideomotor gaze control in a social (vs. non-social) context with respect to temporal (contiguity) and probabilistic (contingency) dynamics could not be found, suggesting parsimonious, common underlying mechanisms. To conclude, the present work contributed to integrate research on gaze interaction, a meaningful domain of social interaction, into general theories of human action control, thereby providing a powerful explanatory framework for mechanisms underlying gaze interaction. Coming back to the initial example with the attention-seeking glances at the party, the following conclusion can therefore be drawn from the present results: Regardless of whether I want to attract the attention of an

attractive person with my gaze or to unlock my smartphone, similar cognitive mechanisms are at stake. For a rather shy person, the idea of unlocking the smartphone with her/his gaze might thus be helpful to use her/his gaze without signs of nervousness at the next party to attract the attention of others.

## REFERENCES

- Abe, K., Owada, K., Ohi, S., & Ohyama, M. (2008). A system for web browsing by eye-gaze input. *Electronics and Communications in Japan*, *91*(5), 11–18. <https://doi.org/10.1002/ecj.10110>
- Adams, N., Witkowski, M., & Spence, R. (2008). The inspection of very large images by eye-gaze control. *Proceedings of the Working Conference on Advanced Visual Interfaces*, 111–118. <https://doi.org/10.1145/1385569.1385589>
- Adams, R. B., Gordon, H. L., Baird, A. A., Ambady, N., & Kleck, R. E. (2003). Effects of gaze on amygdala sensitivity to anger and fear faces. *Science*, *300*(5625), 1536. <https://doi.org/10.1126/science.1082244>
- Adams, R. B., & Kleck, R. E. (2005). Effects of direct and averted gaze on the perception of facially communicated emotion. *Emotion*, *5*(1), 3–11. <https://doi.org/10.1037/1528-3542.5.1.3>
- Akkil, D., University Of Tampere, Isokoski, P., Majaranta, P., Kangas, J., Rantala, J., . . . Špakov, O. (2016). Haptic feedback in eye typing. *Journal of Eye Movement Research*, *9*(1), 1–14. <https://doi.org/10.16910/JEMR.9.1.3>
- Ansorge, U. (2002). Spatial intention–response compatibility. *Acta Psychologica*, *109*(3), 285–299. [https://doi.org/10.1016/S0001-6918\(01\)00062-2](https://doi.org/10.1016/S0001-6918(01)00062-2)
- Asai, K., Osawa, N., Takahashi, H., Sugimoto, Y. Y., Yamazaki, S., Samejima, A., & Tanimae, T. (2000). Eye mark pointer in immersive projection display. *Proceedings of the IEEE Virtual Reality Conference*, 125–132. <https://doi.org/10.1109/VR.2000.840490>
- Badets, A., Koch, I., & Toussaint, L. (2013). Role of an ideomotor mechanism in number processing. *Experimental Psychology*, *60*(1), 34–43. <https://doi.org/10.1027/1618-3169/a000171>
- Bailenson, J. N., Beall, A. C., Loomis, J., Blascovich, J., & Turk, M. (2004). Transformed social interaction: Decoupling representation from behavior and form in collaborative virtual environments. *Presence: Teleoperators and Virtual Environments*, *13*(4), 428–441. <https://doi.org/10.1162/1054746041944803>
- Baron-Cohen, S. (1995). The eye direction detector (EDD) and the shared attention mechanisms (SAM): Two cases for evolutionary psychology. In C. Moore & P. Dunham (Eds.), *Joint attention: Its origins and role in development*. Hillsdale, NJ: Erlbaum.
- Bayle, D. J., Schoendorff, B., Hénaff, M.-A., & Krolak-Salmon, P. (2011). Emotional facial expression detection in the peripheral visual field. *PLoS One*, *6*(6), e21584. <https://doi.org/10.1371/journal.pone.0021584>
- Bayliss, A. P., Murphy, E., Naughtin, C. K., Kritikos, A., Schilbach, L., & Becker, S. I. (2013). "Gaze leading": Initiating simulated joint attention influences eye movements and choice behavior. *Journal of Experimental Psychology: General*, *142*(1), 76–92. <https://doi.org/10.1037/a0029286>
- Bayliss, A. P., Paul, M. A., Cannon, P. R., & Tipper, S. P. (2006). Gaze cuing and affective judgments of objects: I like what you look at. *Psychonomic Bulletin & Review*, *13*(6), 1061–1066. <https://doi.org/10.3758/BF03213926>
- Becchio, C., Manera, V., Sartori, L., Cavallo, A., & Castiello, U. (2012). Grasping intentions: From thought experiments to empirical evidence. *Frontiers in Human Neuroscience*, *6*, 117. <https://doi.org/10.3389/fnhum.2012.00117>

- Beckers, T., Houwer, J. de, & Eelen, P. (2002). Automatic integration of non-perceptual action effect features: The case of the associative affective Simon effect. *Psychological Research*, *66*(3), 166–173. <https://doi.org/10.1007/s00426-002-0090-9>
- Bertelson, P. (1967). The time course of preparation. *The Quarterly Journal of Experimental Psychology*, *19*(3), 272–279. <https://doi.org/10.1080/14640746708400102>
- Bertelson, P., & Tisseyre, F. (1969). The time-course of preparation: Confirmatory results with visual and auditory warning signals. *Acta Psychologica*, *30*, 145–154. [https://doi.org/10.1016/0001-6918\(69\)90047-X](https://doi.org/10.1016/0001-6918(69)90047-X)
- Birmingham, E., Ristic, J., & Kingstone, A. (2012). Investigating social attention: A case for increasing stimulus complexity in the laboratory. In J. A. Burack, J. T. Enns, & N. A. Fox (Eds.), *Cognitive Neuroscience, Development, and Psychopathology* (pp. 251–276). Oxford, England: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195315455.003.0010>
- Böckler, A., van der Wel, R. P. R. D., & Welsh, T. N. (2014). Catching eyes: Effects of social and nonsocial cues on attention capture. *Psychological Science*, *25*(3), 720–727. <https://doi.org/10.1177/0956797613516147>
- Böckler, A., van der Wel, R. P. R. D., & Welsh, T. N. (2015). Eyes only? Perceiving eye contact is neither sufficient nor necessary for attentional capture by face direction. *Acta Psychologica*, *160*, 134–140. <https://doi.org/10.1016/j.actpsy.2015.07.009>
- Bompas, A., Hedge, C., & Sumner, P. (2017). Speeded saccadic and manual visuo-motor decisions: Distinct processes but same principles. *Cognitive Psychology*, *94*, 26–52. <https://doi.org/10.1016/j.cogpsych.2017.02.002>
- Boyer, T. W., & Wang, M. (2018). Direct gaze, eye movements, and covert and overt social attention processes. *Attention, Perception & Psychophysics*, *80*(7), 1654–1659. <https://doi.org/10.3758/s13414-018-1590-z>
- Brass, M., & Haggard, P. (2008). The what, when, whether model of intentional action. *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *14*(4), 319–325. <https://doi.org/10.1177/1073858408317417>
- Bruce, V., & Langton, S. (1994). The use of pigmentation and shading information in recognising the sex and identities of faces. *Perception*, *23*(7), 803–822. <https://doi.org/10.1068/p230803>
- Bucker, B., Silvis, J. D., Donk, M., & Theeuwes, J. (2015). Reward modulates oculomotor competition between differently valued stimuli. *Vision Research*, *108*, 103–112. <https://doi.org/10.1016/j.visres.2015.01.020>
- Bunlon, F., Marshall, P. J., Quandt, L. C., & Bouquet, C. A. (2015). Influence of action-effect associations acquired by ideomotor learning on imitation. *PloS One*, *10*(3), e0121617. <https://doi.org/10.1371/journal.pone.0121617>
- Camus, T., Hommel, B., Brunel, L., & Brouillet, T. (2018). From anticipation to integration: The role of integrated action-effects in building sensorimotor contingencies. *Psychonomic Bulletin & Review*, *25*(3), 1059–1065. <https://doi.org/10.3758/s13423-017-1308-6>
- Carrasco, M., Evert, D. L., Chang, I., & Katz, S. M. (1995). The eccentricity effect: Target eccentricity affects performance on conjunction searches. *Perception & Psychophysics*, *57*(8), 1241–1261. <https://doi.org/10.3758/bf03208380>
- Carrasco, M., McElree, B., Denisova, K., & Giordano, A. M. (2003). Speed of visual processing increases with eccentricity. *Nature Neuroscience*, *6*(7), 699–700. <https://doi.org/10.1038/nn1079>
- Catmur, C., & Heyes, C. (2013). Is it what you do, or when you do it? The roles of contingency and similarity in pro-social effects of imitation. *Cognitive Science*, *37*(8), 1541–1552. <https://doi.org/10.1111/cogs.12071>
- Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Cox, A., & Drew, A. (2000). Testing joint attention, imitation, and play as infancy precursors to language and theory of mind. *Cognitive Development*, *15*(4), 481–498. [https://doi.org/10.1016/S0885-2014\(01\)00037-5](https://doi.org/10.1016/S0885-2014(01)00037-5)
- Chen, J., & Proctor, R. W. (2013). Response-effect compatibility defines the natural scrolling direction. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *55*(6), 1112–1129. <https://doi.org/10.1177/0018720813482329>

- Chiappe, D., Brown, A., & Dow, B. (2004). Cheaters are looked at longer and remembered better than cooperators in social exchange situations. *Evolutionary Psychology*, 2(1), 108-120. <https://doi.org/10.1177/147470490400200117>
- Corcoran, P., Nanu, F., Petrescu, S., & Bigioi, P. (2012). Real-time eye gaze tracking for gaming design and consumer electronics systems. *IEEE Transactions on Consumer Electronics*, 58(2), 347-355. <https://doi.org/10.1109/TCE.2012.6227433>
- Dalmaso, M., Edwards, S. G., & Bayliss, A. P. (2016). Re-encountering individuals who previously engaged in joint gaze modulates subsequent gaze cueing. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 42(2), 271-284. <https://doi.org/10.1037/xlm0000159>
- Deaner, R. O., Shepherd, S. V., & Platt, M. L. (2007). Familiarity accentuates gaze cuing in women but not men. *Biology Letters*, 3(1), 65-68. <https://doi.org/10.1098/rsbl.2006.0564>
- Dignath, D., & Janczyk, M. (2017). Anticipation of delayed action-effects: Learning when an effect occurs, without knowing what this effect will be. *Psychological Research*, 81(5), 1072-1083. <https://doi.org/10.1007/s00426-016-0797-7>
- Dignath, D., Lotze-Hermes, P., Farmer, H., & Pfister, R. (2018). Contingency and contiguity of imitative behaviour affect social affiliation. *Psychological Research*, 82(4), 819-831. <https://doi.org/10.1007/s00426-017-0854-x>
- Dignath, D., Pfister, R., Eder, A. B., Kiesel, A., & Kunde, W. (2014). Representing the hyphen in action-effect associations: Automatic acquisition and bidirectional retrieval of action-effect intervals. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 40(6), 1701-1712. <https://doi.org/10.1037/xlm0000022>
- Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition*, 6(5), 509-540. <https://doi.org/10.1080/135062899394920>
- Drost, U. C., Rieger, M., Brass, M., Gunter, T. C., & Prinz, W. (2005). Action-effect coupling in pianists. *Psychological Research*, 69(4), 233-241. <https://doi.org/10.1007/s00426-004-0175-8>
- Dunne, S., Ellison, A., & Smith, D. T. (2015). Rewards modulate saccade latency but not exogenous spatial attention. *Frontiers in Psychology*, 6, 1080. <https://doi.org/10.3389/fpsyg.2015.01080>
- Edwards, S. G., Stephenson, L. J., Dalmaso, M., & Bayliss, A. P. (2015). Social orienting in gaze leading: A mechanism for shared attention. *Proceedings of the Royal Society B*, 282(1812), 20151141. <https://doi.org/10.1098/rspb.2015.1141>
- Elsner, B., & Hommel, B. (2001). Effect anticipation and action control. *Journal of Experimental Psychology. Human Perception and Performance*, 27(1), 229-240. <https://doi.org/10.1037/0096-1523.27.1.229>
- Elsner, B., & Hommel, B. (2004). Contiguity and contingency in action-effect learning. *Psychological Research*, 68(2-3), 138-154. <https://doi.org/10.1007/s00426-003-0151-8>
- Emery, N. J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience & Biobehavioral Reviews*, 24(6), 581-604. [https://doi.org/10.1016/S0149-7634\(00\)00025-7](https://doi.org/10.1016/S0149-7634(00)00025-7)
- Enns, J. T., Austen, E. L., Di Lollo, V., Rauschenberger, R., & Yantis, S. (2001). New objects dominate luminance transients in setting attentional priority. *Journal of Experimental Psychology. Human Perception and Performance*, 27(6), 1287-1302. <https://doi.org/10.1037//0096-1523.27.6.1287>
- Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the United States of America*, 99(14), 9602-9605. <https://doi.org/10.1073/pnas.152159999>
- Farroni, T., Johnson, M. H., Brockbank, M., & Simion, F. (2000). Infants' use of gaze direction to cue attention: The importance of perceived motion. *Visual Cognition*, 7(6), 705-718. <https://doi.org/10.1080/13506280050144399>
- Farroni, T., Johnson, M. H., & Csibra, G. (2004). Mechanisms of eye gaze perception during infancy. *Journal of Cognitive Neuroscience*, 16(8), 1320-1326. <https://doi.org/10.1162/0898929042304787>

- Farroni, T., Massaccesi, S., Pividori, D., & Johnson, M. H. (2004). Gaze following in newborns. *Infancy*, 5(1), 39–60. [https://doi.org/10.1207/s15327078in0501\\_2](https://doi.org/10.1207/s15327078in0501_2)
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Findlay, J. M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral and Brain Sciences*, 22(4), 661–74; discussion 674–721.
- Fitts, P. M., & Seeger, C. M. (1953). S-R compatibility: spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology. General*, 46(3), 199–210. <https://doi.org/10.1037/h0062827>
- Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze: Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review*, 5(3), 490–495. <https://doi.org/10.3758/BF03208827>
- Friesen, C. K., & Kingstone, A. (2003). Covert and overt orienting to gaze direction cues and the effects of fixation offset. *Neuroreport*, 14(3), 489–493. <https://doi.org/10.1097/00001756-200303030-00039>
- Friesen, C. K., Ristic, J., & Kingstone, A. (2004). Attentional effects of counterpredictive gaze and arrow cues. *Journal of Experimental Psychology. Human Perception and Performance*, 30(2), 319–329. <https://doi.org/10.1037/0096-1523.30.2.319>
- Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin*, 133(4), 694–724. <https://doi.org/10.1037/0033-2909.133.4.694>
- Frith, C. D., & Frith, U. (2008). Implicit and explicit processes in social cognition. *Neuron*, 60(3), 503–510. <https://doi.org/10.1016/j.neuron.2008.10.032>
- Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends in Cognitive Sciences*, 4(1), 14–21. [https://doi.org/10.1016/S1364-6613\(99\)01417-5](https://doi.org/10.1016/S1364-6613(99)01417-5)
- Gallup, A. C., Hale, J. J., Sumpter, D. J. T., Garnier, S., Kacelnik, A., Krebs, J. R., & Couzin, I. D. (2012). Visual attention and the acquisition of information in human crowds. *Proceedings of the National Academy of Sciences of the United States of America*, 109(19), 7245–7250. <https://doi.org/10.1073/pnas.1116141109>
- Gamer, M., Hecht, H., Seipp, N., & Hiller, W. (2011). Who is looking at me? The cone of gaze widens in social phobia. *Cognition & Emotion*, 25(4), 756–764. <https://doi.org/10.1080/02699931.2010.503117>
- Gobel, M. S., Kim, H. S., & Richardson, D. C. (2015). The dual function of social gaze. *Cognition*, 136, 359–364. <https://doi.org/10.1016/j.cognition.2014.11.040>
- Graefe, T. M., & Vaughan, J. (1978). Saccadic and manual reaction times to stimuli initiated by eye or finger movements. *Bulletin of the Psychonomic Society*, 11(2), 97–99. <https://doi.org/10.3758/BF03336776>
- Grafton, S. T., Arbib, M. A., Fadiga, L., & Rizzolatti, G. (1996). Localization of grasp representations in humans by positron emission tomography: 2. Observation compared with imagination. *Experimental Brain Research*, 112(1), 103–111. <https://doi.org/10.1007/bf00227183>
- Greenwald, A. G. (1970). Sensory feedback mechanisms in performance control: With special reference to the ideo-motor mechanism. *Psychological Review*, 77(2), 73–99. <https://doi.org/10.1037/h0028689>
- Grosbras, M.-H., Laird, A. R., & Paus, T. (2005). Cortical regions involved in eye movements, shifts of attention, and gaze perception. *Human Brain Mapping*, 25(1), 140–154. <https://doi.org/10.1002/hbm.20145>
- Grünau, M. von, & Anston, C. (1995). The detection of gaze direction: A stare-in-the-crowd effect. *Perception*, 24(11), 1297–1313. <https://doi.org/10.1068/p241297>
- Haering, C., & Kiesel, A. (2012). Time in action contexts: Learning when an action effect occurs. *Psychological Research*, 76(3), 336–344. <https://doi.org/10.1007/s00426-011-0341-8>
- Haggard, P. (2005). Conscious intention and motor cognition. *Trends in Cognitive Sciences*, 9(6), 290–295. <https://doi.org/10.1016/j.tics.2005.04.012>

- Haley, K. J., & Fessler, D. M. T. (2005). Nobody's watching?: Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior*, 26(3), 245–256. <https://doi.org/10.1016/j.evolhumbehav.2005.01.002>
- Hallett, P. E. (1978). Primary and secondary saccades to goals defined by instructions. *Vision Research*, 18(10), 1279–1296. [https://doi.org/10.1016/0042-6989\(78\)90218-3](https://doi.org/10.1016/0042-6989(78)90218-3)
- Halwani, Y., Salcudean, S. E., Lessoway, V. A., & Fels, S. S. (2017). Enhancing zoom and pan in ultrasound machines with a multimodal gaze-based interface. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 1648–1654. <https://doi.org/10.1145/3027063.3053174>
- Harleß, E. (1861). Der Apparat des Willens [The apparatus of will]. *Zeitschrift Für Philosophie Und Philosophische Kritik*, 38, 50–73.
- Hayward, D. A., Voorhies, W., Morris, J. L., Capozzi, F., & Ristic, J. (2017). Staring reality in the face: A comparison of social attention across laboratory and real world measures suggests little common ground. *Canadian Journal of Experimental Psychology*, 71(3), 212–225. <https://doi.org/10.1037/cep0000117>
- Heider, F. (1977). *Psychologie der interpersonalen Beziehungen [The psychology of interpersonal relations]*. Stuttgart, Germany: Klett.
- Herbart, J. F. (1825). *Psychologie als Wissenschaft neu gegründet auf Erfahrung, Metaphysik, und Mathematik. Zweiter analytischer Teil [Psychology as a science newly founded on experience, metaphysics, and mathematics. Second analytical part]*. Königsberg, Germany: Unzer.
- Herbort, O., Koning, A., van Uem, J., & Meulenbroek, R. G. (2012). The end-state comfort effect facilitates joint action. *Acta Psychologica*, 139(3), 404–416. <https://doi.org/10.1016/j.actpsy.2012.01.001>
- Herbort, O., & Kunde, W. (2016). Spatial (mis-)interpretation of pointing gestures to distal referents. *Journal of Experimental Psychology. Human Perception and Performance*, 42(1), 78–89. <https://doi.org/10.1037/xhp0000126>
- Herbort, O., & Kunde, W. (2018). How to point and to interpret pointing gestures? Instructions can reduce pointer-observer misunderstandings. *Psychological Research*, 82(2), 395–406. <https://doi.org/10.1007/s00426-016-0824-8>
- Herwig, A. (2015). Linking perception and action by structure or process? Toward an integrative perspective. *Neuroscience and Biobehavioral Reviews*, 52, 105–116. <https://doi.org/10.1016/j.neubiorev.2015.02.013>
- Herwig, A., Beisert, M., & Prinz, W. (2013). Action science emerging: Introduction and leitmotifs. In W. Prinz, M. Beisert, & A. Herwig (Eds.), *Action science: Foundations of an emerging discipline* (1–33). Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/9780262018555.003.0001>
- Herwig, A., & Horstmann, G. (2011). Action-effect associations revealed by eye movements. *Psychonomic Bulletin & Review*, 18(3), 531–537. <https://doi.org/10.3758/s13423-011-0063-3>
- Herwig, A., Prinz, W., & Waszak, F. (2007). Two modes of sensorimotor integration in intention-based and stimulus-based actions. *Quarterly Journal of Experimental Psychology*, 60(11), 1540–1554. <https://doi.org/10.1080/17470210601119134>
- Herwig, A., & Schneider, W. X. (2014). Predicting object features across saccades: Evidence from object recognition and visual search. *Journal of Experimental Psychology. General*, 143(5), 1903–1922. <https://doi.org/10.1037/a0036781>
- Herwig, A., & Waszak, F. (2009). Intention and attention in ideomotor learning. *Quarterly Journal of Experimental Psychology*, 62(2), 219–227. <https://doi.org/10.1080/17470210802373290>
- Herwig, A., & Waszak, F. (2012). Action-effect bindings and ideomotor learning in intention- and stimulus-based actions. *Frontiers in Psychology*, 3, 444. <https://doi.org/10.3389/fpsyg.2012.00444>
- Herwig, A., Weiss, K., & Schneider, W. X. (2015). When circles become triangular: How transsaccadic predictions shape the perception of shape. *Annals of the New York Academy of Sciences*, 1339, 97–105. <https://doi.org/10.1111/nyas.12672>
- Heyes, C. (2001). Causes and consequences of imitation. *Trends in Cognitive Sciences*, 5(6), 253–261. [https://doi.org/10.1016/S1364-6613\(00\)01661-2](https://doi.org/10.1016/S1364-6613(00)01661-2)

- Heyes, C. (2011). Automatic imitation. *Psychological Bulletin*, *137*(3), 463–483. <https://doi.org/10.1037/a0022288>
- Heyes, C., Bird, G., Johnson, H., & Haggard, P. (2005). Experience modulates automatic imitation. *Cognitive Brain Research*, *22*(2), 233–240. <https://doi.org/10.1016/j.cogbrainres.2004.09.009>
- Hickey, C., & van Zoest, W. (2012). Reward creates oculomotor salience. *Current Biology : CB*, *22*(7), R219–20. <https://doi.org/10.1016/j.cub.2012.02.007>
- Hietanen, J. K., & Leppänen, J. M. (2003). Does facial expression affect attention orienting by gaze direction cues? *Journal of Experimental Psychology. Human Perception and Performance*, *29*(6), 1228–1243. <https://doi.org/10.1037/0096-1523.29.6.1228>
- Hoffmann, J., Lenhard, A., Sebald, A., & Pfister, R. (2009). Movements or targets: What makes an action in action-effect learning? *Quarterly Journal of Experimental Psychology*, *62*(12), 2433–2449. <https://doi.org/10.1080/17470210902922079>
- Hommel, B. (1993). Inverting the Simon effect by intention. *Psychological Research*, *55*(4), 270–279. <https://doi.org/10.1007/BF00419687>
- Hommel, B. (1996). The cognitive representation of action: Automatic integration of perceived action effects. *Psychological Research*, *59*(3), 176–186. <https://doi.org/10.1007/BF00425832>
- Hommel, B. (2009). Action control according to TEC (theory of event coding). *Psychological Research*, *73*(4), 512–526. <https://doi.org/10.1007/s00426-009-0234-2>
- Hommel, B. (2013). Ideomotor action control: On the perceptual grounding of voluntary actions and agents. In W. Prinz, M. Beisert, & A. Herwig (Eds.), *Action science: Foundations of an emerging discipline* (pp. 112–136). Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/9780262018555.003.0005>
- Hommel, B. (2019). Theory of Event Coding (TEC) V2.0: Representing and controlling perception and action. *Attention, Perception & Psychophysics*, *81*(7), 2139–2154. <https://doi.org/10.3758/s13414-019-01779-4>
- Hommel, B., Alonso, D., & Fuentes, L. (2003). Acquisition and generalization of action effects. *Visual Cognition*, *10*(8), 965–986. <https://doi.org/10.1080/13506280344000176>
- Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, *24*(5), 849–878. <https://doi.org/10.1017/S0140525X01000103>
- Hood, B. M., Macrae, C. N., Cole-Davies, V., & Dias, M. (2003). Eye remember you: the effects of gaze direction on face recognition in children and adults. *Developmental Science*, *6*(1), 67–71. <https://doi.org/10.1111/1467-7687.00256>
- Hornof, A. J., & Cavender, A. (2005). EyeDraw. *Proceedings of the Conference on Human Factors in Computing Systems*, 161–170. <https://doi.org/10.1145/1054972.1054995>
- Hortensius, R., & Gelder, B. de (2018). From empathy to apathy: The bystander effect revisited. *Current Directions in Psychological Science*, *27*(4), 249–256. <https://doi.org/10.1177/0963721417749653>
- Hudson, M., Nicholson, T., Simpson, W. A., Ellis, R., & Bach, P. (2015). One step ahead: The perceived kinematics of others' actions are biased toward expected goals. *Journal of Experimental Psychology. General*, *145*(1), 1–7. <https://doi.org/10.1037/xge0000126>
- Huestegge, L. (2011). The role of saccades in multitasking: Towards an output-related view of eye movements. *Psychological Research*, *75*(6), 452–465. <https://doi.org/10.1007/s00426-011-0352-5>
- Huestegge, L., Herbort, O., Gosch, N., Kunde, W., & Pieczykolan, A. (2019). Free-choice saccades and their underlying determinants: Explorations of high-level voluntary oculomotor control. *Journal of Vision*, *19*(3), 14. <https://doi.org/10.1167/19.3.14>
- Huestegge, L., & Koch, I. (2009). Dual-task crosstalk between saccades and manual responses. *Journal of Experimental Psychology. Human Perception and Performance*, *35*(2), 352–362. <https://doi.org/10.1037/a0013897>
- Huestegge, L., & Koch, I. (2010a). Crossmodal action selection: Evidence from dual-task compatibility. *Memory & Cognition*, *38*(4), 493–501. <https://doi.org/10.3758/MC.38.4.493>



- Huestegge, L., & Koch, I. (2010b). Fixation disengagement enhances peripheral perceptual processing: Evidence for a perceptual gap effect. *Experimental Brain Research*, 201(4), 631–640. <https://doi.org/10.1007/s00221-009-2080-2>
- Huestegge, L., & Koch, I. (2013). Constraints in task-set control: Modality dominance patterns among effector systems. *Journal of Experimental Psychology. General*, 142(3), 633–637. <https://doi.org/10.1037/a0030156>
- Huestegge, L., & Kreutzfeldt, M. (2012). Action effects in saccade control. *Psychonomic Bulletin & Review*, 19(2), 198–203. <https://doi.org/10.3758/s13423-011-0215-5>
- Huestegge, L., Pieczykolan, A., & Koch, I. (2014). Talking while looking: On the encapsulation of output system representations. *Cognitive Psychology*, 73, 72–91. <https://doi.org/10.1016/j.cogpsych.2014.06.001>
- James, W. (1890). *The principles of psychology*. New York, NY: Henry Holt.
- Janczyk, M., Yamaguchi, M., Proctor, R. W., & Pfister, R. (2015). Response-effect compatibility with complex actions: The case of wheel rotations. *Attention, Perception & Psychophysics*, 77(3), 930–940. <https://doi.org/10.3758/s13414-014-0828-7>
- Jones, B. C., DeBruine, L. M., Little, A. C., Conway, C. A., & Feinberg, D. R. (2006). Integrating gaze direction and expression in preferences for attractive faces. *Psychological Science*, 17(7), 588–591. <https://doi.org/10.1111/j.1467-9280.2006.01749.x>
- Jones, B. C., DeBruine, L. M., Main, J. C., Little, A. C., Welling, L. L. M., Feinberg, D. R., & Tideman, B. P. (2010). Facial cues of dominance modulate the short-term gaze-cuing effect in human observers. *Proceedings of the Royal Society B*, 277, 617–624. <https://doi.org/10.1098/rspb.2009.1575>
- Joyce, K., Schenke, K., Bayliss, A., & Bach, P. (2016). Looking ahead: Anticipatory cueing of attention to objects others will look at. *Cognitive Neuroscience*, 7(1-4), 74–81. <https://doi.org/10.1080/17588928.2015.1053443>
- Jüttner, M., & Rentschler, I. (2000). Scale-invariant superiority of foveal vision in perceptual categorization. *The European Journal of Neuroscience*, 12(1), 353–359. <https://doi.org/10.1046/j.1460-9568.2000.00907.x>
- Kano, F., & Call, J. (2014). Cross-species variation in gaze following and conspecific preference among great apes, human infants and adults. *Animal Behaviour*, 91, 137–150. <https://doi.org/10.1016/j.anbehav.2014.03.011>
- Keller, P. E., & Koch, I. (2006). Exogenous and endogenous response priming with auditory stimuli. *Advances in Cognitive Psychology*, 2(4), 269–276. <https://doi.org/10.2478/v10053-008-0061-9>
- Keller, P. E., & Koch, I. (2008). Action planning in sequential skills: Relations to music performance. *Quarterly Journal of Experimental Psychology*, 61(2), 275–291. <https://doi.org/10.1080/17470210601160864>
- Keller, P. E., Wascher, E., Prinz, W., Waszak, F., Koch, I., & Rosenbaum, D. A. (2006). Differences between intention-based and stimulus-based actions. *Journal of Psychophysiology*, 20(1), 9–20. <https://doi.org/10.1027/0269-8803.20.1.9>
- Kiesel, A., & Hoffmann, J. (2004). Variable action effects: Response control by context-specific effect anticipations. *Psychological Research*, 68(2-3), 155–162. <https://doi.org/10.1007/s00426-003-0152-7>
- Kim, D., & Hommel, B. (2019). Social cognition 2.0: Toward mechanistic theorizing. *Frontiers in Psychology*, 10, 97. <https://doi.org/10.3389/fpsyg.2019.02643>
- Kingstone, A., Friesen, C. K., & Gazzaniga, M. S. (2000). Reflexive joint attention depends on lateralized cortical connections. *Psychological Science*, 11(2), 159–166. <https://doi.org/10.1111/1467-9280.00232>
- Kingstone, A., & Klein, R. M. (1993). Visual offsets facilitate saccadic latency: Does predisengagement of visuospatial attention mediate this gap effect? *Journal of Experimental Psychology. Human Perception and Performance*, 19(6), 1251–1265. <https://doi.org/10.1037/0096-1523.19.6.1251>

- Knoblich, G., & Jordan, J. S. (2003). Action coordination in groups and individuals: Learning anticipatory control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(5), 1006–1016. <https://doi.org/10.1037/0278-7393.29.5.1006>
- Koch, I., & Kunde, W. (2002). Verbal response-effect compatibility. *Memory & Cognition*, 30(8), 1297–1303. <https://doi.org/10.3758/BF03213411>
- Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus-response compatibility--A model and taxonomy. *Psychological Review*, 97(2), 253–270. <https://doi.org/10.1037/0033-295X.97.2.253>
- Kuhn, G., & Benson, V. (2007). The influence of eye-gaze and arrow pointing distractor cues on voluntary eye movements. *Perception & Psychophysics*, 69(6), 966–971. <https://doi.org/10.3758/BF03193934>
- Kuhn, G., & Kingstone, A. (2009). Look away! Eyes and arrows engage oculomotor responses automatically. *Attention, Perception & Psychophysics*, 71(2), 314–327. <https://doi.org/10.3758/APP.71.2.314>
- Kühn, S., & Brass, M. (2010). The cognitive representation of intending not to act: Evidence for specific non-action-effect binding. *Cognition*, 117(1), 9–16. <https://doi.org/10.1016/j.cognition.2010.06.006>
- Kühn, S., Elsner, B., Prinz, W., & Brass, M. (2009). Busy doing nothing: Evidence for nonaction-effect binding. *Psychonomic Bulletin & Review*, 16(3), 542–549. <https://doi.org/10.3758/PBR.16.3.542>
- Kunde, W. (2001). Response-effect compatibility in manual choice reaction tasks. *Journal of Experimental Psychology: Human Perception and Performance*, 27(2), 387–394. <https://doi.org/10.1037/0096-1523.27.2.387>
- Kunde, W. (2003). Temporal response-effect compatibility. *Psychological Research*, 67(3), 153–159. <https://doi.org/10.1007/s00426-002-0114-5>
- Kunde, W. (2004). Response priming by supraliminal and subliminal action effects. *Psychological Research*, 68(2-3), 91–96. <https://doi.org/10.1007/s00426-003-0147-4>
- Kunde, W., Koch, I., & Hoffmann, J. (2004). Anticipated action effects affect the selection, initiation, and execution of actions. *The Quarterly Journal of Experimental Psychology*, 57(1), 87–106. <https://doi.org/10.1080/02724980343000143>
- Kunde, W., Lozo, L., & Neumann, R. (2011). Effect-based control of facial expressions: Evidence from action-effect compatibility. *Psychonomic Bulletin & Review*, 18(4), 820–826. <https://doi.org/10.3758/s13423-011-0093-x>
- Kunde, W., & Weigelt, M. (2005). Goal congruency in bimanual object manipulation. *Journal of Experimental Psychology: Human Perception and Performance*, 31(1), 145–156. <https://doi.org/10.1037/0096-1523.31.1.145>
- Kunde, W., Weller, L., & Pfister, R. (2018). Sociomotor action control. *Psychonomic Bulletin & Review*, 25(3), 917–931. <https://doi.org/10.3758/s13423-017-1316-6>
- Lachat, F., Conty, L., Hugueville, L., & George, N. (2012). Gaze cueing effect in a face-to-face situation. *Journal of Nonverbal Behavior*, 36(3), 177–190. <https://doi.org/10.1007/s10919-012-0133-x>
- Laidlaw, K. E. W., Foulsham, T., Kuhn, G., & Kingstone, A. (2011). Potential social interactions are important to social attention. *Proceedings of the National Academy of Sciences of the United States of America*, 108(14), 5548–5553. <https://doi.org/10.1073/pnas.1017022108>
- Land, M., Mennie, N., & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception*, 28(11), 1311–1328. <https://doi.org/10.1068/p2935>
- Land, M. F. (2006). Eye movements and the control of actions in everyday life. *Progress in Retinal and Eye Research*, 25(3), 296–324. <https://doi.org/10.1016/j.preteyeres.2006.01.002>
- Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, 41(25-26), 3559–3565. [https://doi.org/10.1016/S0042-6989\(01\)00102-X](https://doi.org/10.1016/S0042-6989(01)00102-X)
- Land, M. F., & Tatler, B. W. (2009). *Looking and acting: Vision and eye movements in natural behaviour*. Oxford, England: Oxford University Press.

- Langton, S. R.H., & Bruce, V. (1999). Reflexive visual orienting in response to the social attention of others. *Visual Cognition*, 6(5), 541–567. <https://doi.org/10.1080/135062899394939>
- Lauwereyns, J., Watanabe, K., Coe, B., & Hikosaka, O. (2002). A neural correlate of response bias in monkey caudate nucleus. *Nature*, 418(6896), 413–417. <https://doi.org/10.1038/nature00892>
- Lelonkiewicz, J. R., & Gambi, C. (2017). Spontaneous adaptation explains why people act faster when being imitated. *Psychonomic Bulletin & Review*, 24(3), 842–848. <https://doi.org/10.3758/s13423-016-1141-3>
- Lelonkiewicz, J. R., Gambi, C., Weller, L., & Pfister, R. (2020). Action-effect anticipation and temporal adaptation in social interactions. *Journal of Experimental Psychology: Human Perception and Performance*, 46(4), 335–349. <https://doi.org/10.1037/xhp0000717>
- Liversedge, S., Gilchrist, I., & Everling, S. (2011). *The Oxford Handbook of Eye Movements*. Oxford, England: Oxford University Press.
- Los, S. A., & Schut, M. L. J. (2008). The effective time course of preparation. *Cognitive Psychology*, 57(1), 20–55. <https://doi.org/10.1016/j.cogpsych.2007.11.001>
- Lotze, R. H. (1852). *Medicinische Psychologie oder Physiologie der Seele*. Leipzig, Germany: Weidmann'sche Buchhandlung.
- Macrae, C. N., Hood, B. M., Milne, A. B., Rowe, A. C., & Mason, M. F. (2002). Are you looking at me? Eye gaze and person perception. *Psychological Science*, 13(5), 460–464. <https://doi.org/10.1111/1467-9280.00481>
- Majaranta, P., Rähkä, K.-J., Hyrskykari, A., & Špakov, O. (2019). Eye movements and human-computer interaction. In C. Klein & U. Ettinger (Eds.), *Eye movement research: An introduction to its scientific foundations and applications* (pp. 971–1015). Cham, Switzerland: Springer. [https://doi.org/10.1007/978-3-030-20085-5\\_23](https://doi.org/10.1007/978-3-030-20085-5_23)
- Mann, S., Vrij, A., Leal, S., Granhag, P. A., Warmelink, L., & Forrester, D. (2012). Windows to the soul? Deliberate eye contact as a cue to deceit. *Journal of Nonverbal Behavior*, 36(3), 205–215. <https://doi.org/10.1007/s10919-012-0132-y>
- Mansfield, E., Farroni, T., & Johnson, M. (2003). Does gaze perception facilitate overt orienting? *Visual Cognition*, 10(1), 7–14. <https://doi.org/10.1080/713756671>
- Massen, C. (2004). Parallel programming of exogenous and endogenous components in the antisaccade task. *The Quarterly Journal of Experimental Psychology*, 57(3), 475–498. <https://doi.org/10.1080/02724980343000341>
- Matarić, M. J. (2017). Socially assistive robotics: Human augmentation versus automation. *Science Robotics*, 2(4), eaam5410. <https://doi.org/10.1126/scirobotics.aam5410>
- McCoy, B., & Theeuwes, J. (2016). Effects of reward on oculomotor control. *Journal of Neurophysiology*, 116(5), 2453–2466. <https://doi.org/10.1152/jn.00498.2016>
- Memelink, J., & Hommel, B. (2013). Intentional weighting: A basic principle in cognitive control. *Psychological Research*, 77(3), 249–259. <https://doi.org/10.1007/s00426-012-0435-y>
- Milstein, D. M., & Dorris, M. C. (2007). The influence of expected value on saccadic preparation. *The Journal of Neuroscience*, 27(18), 4810–4818. <https://doi.org/10.1523/JNEUROSCI.0577-07.2007>
- Mitchell, J. P. (2009). Social psychology as a natural kind. *Trends in Cognitive Sciences*, 13(6), 246–251. <https://doi.org/10.1016/j.tics.2009.03.008>
- Moore, C. (2008). The development of gaze following. *Child Development Perspectives*, 2(2), 66–70. <https://doi.org/10.1111/j.1750-8606.2008.00052.x>
- Morales, M., Mundy, P., & Rojas, J. (1998). Following the direction of gaze and language development in 6-month-olds. *Infant Behavior and Development*, 21(2), 373–377. [https://doi.org/10.1016/S0163-6383\(98\)90014-5](https://doi.org/10.1016/S0163-6383(98)90014-5)
- Mott, M. E., Williams, S., Wobbrock, J. O., & Morris, M. R. (2017). Improving dwell-based gaze typing with dynamic, cascading dwell times. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2558–2570. <https://doi.org/10.1145/3025453.3025517>

- Von Mühlénen, A., Rempel, M. I., & Enns, J. T. (2005). Unique temporal change is the key to attentional capture. *Psychological Science*, *16*(12), 979–986. <https://doi.org/10.1111/j.1467-9280.2005.01647.x>
- Müller, R. (2016). Does the anticipation of compatible partner reactions facilitate action planning in joint tasks? *Psychological Research*, *80*(4), 464–486. <https://doi.org/10.1007/s00426-015-0670-0>
- Mundy, P. (1995). Joint attention and social-emotional approach behavior in children with autism. *Development and Psychopathology*, *7*(1), 63–82. <https://doi.org/10.1017/S0954579400006349>
- Mundy, P., & Newell, L. (2007). Attention, joint attention, and social cognition. *Current Directions in Psychological Science*, *16*(5), 269–274. <https://doi.org/10.1111/j.1467-8721.2007.00518.x>
- Näsänen, R., Ojanpää, H., & Kojo, I. (2001). Effect of stimulus contrast on performance and eye movements in visual search. *Vision Research*, *41*(14), 1817–1824. [https://doi.org/10.1016/S0042-6989\(01\)00056-6](https://doi.org/10.1016/S0042-6989(01)00056-6)
- Paelecke, M., & Kunde, W. (2007). Action-effect codes in and before the central bottleneck: Evidence from the psychological refractory period paradigm. *Journal of Experimental Psychology. Human Perception and Performance*, *33*(3), 627–644. <https://doi.org/10.1037/0096-1523.33.3.627>
- Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, *87*(6), 532–552. <https://doi.org/10.1037/0033-295X.87.6.532>
- Pecenka, N., & Keller, P. E. (2011). The role of temporal prediction abilities in interpersonal sensorimotor synchronization. *Experimental Brain Research*, *211*(3-4), 505–515. <https://doi.org/10.1007/s00221-011-2616-0>
- Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, *91*(1), 176–180. <https://doi.org/10.1007/bf00230027>
- Pfeiffer, U. J., Vogeley, K., & Schilbach, L. (2013). From gaze cueing to dual eye-tracking: Novel approaches to investigate the neural correlates of gaze in social interaction. *Neuroscience and Biobehavioral Reviews*, *37*(10 Pt 2), 2516–2528. <https://doi.org/10.1016/j.neubiorev.2013.07.017>
- Pfeuffer, C. U., Kiesel, A., & Huestegge, L. (2016). A look into the future: Spontaneous anticipatory saccades reflect processes of anticipatory action control. *Journal of Experimental Psychology. General*, *145*(11), 1530–1547. <https://doi.org/10.1037/xge0000224>
- Pfister, R., Melcher, T., Kiesel, A., Dechent, P., & Gruber, O. (2014). Neural correlates of ideomotor effect anticipations. *Neuroscience*, *259*, 164–171. <https://doi.org/10.1016/j.neurosci.2013.11.061>
- Pfister, R. (2019). Effect-based action control with body-related effects: Implications for empirical approaches to ideomotor action control. *Psychological Review*, *126*(1), 153–161. <https://doi.org/10.1037/rev0000140>
- Pfister, R., Dignath, D., Hommel, B., & Kunde, W. (2013). It takes two to imitate: Anticipation and imitation in social interaction. *Psychological Science*, *24*(10), 2117–2121. <https://doi.org/10.1177/0956797613489139>
- Pfister, R., Janczyk, M., Gressmann, M., Fournier, L. R., & Kunde, W. (2014). Good vibrations? Vibrotactile self-stimulation reveals anticipation of body-related action effects in motor control. *Experimental Brain Research*, *232*(3), 847–854. <https://doi.org/10.1007/s00221-013-3796-6>
- Pfister, R., Kiesel, A., & Hoffmann, J. (2011). Learning at any rate: Action-effect learning for stimulus-based actions. *Psychological Research*, *75*(1), 61–65. <https://doi.org/10.1007/s00426-010-0288-1>
- Pfister, R., Kiesel, A., & Melcher, T. (2010). Adaptive control of ideomotor effect anticipations. *Acta Psychologica*, *135*(3), 316–322. <https://doi.org/10.1016/j.actpsy.2010.08.006>
- Pfister, R., & Kunde, W. (2013). Dissecting the response in response-effect compatibility. *Experimental Brain Research*, *224*(4), 647–655. <https://doi.org/10.1007/s00221-012-3343-x>
- Pfister, R., Obhi, S. S., Rieger, M., & Wenke, D. (2014). Action and perception in social contexts: Intentional binding for social action effects. *Frontiers in Human Neuroscience*, *8*, 667. <https://doi.org/10.3389/fnhum.2014.00667>

- Pfister, R., Pfeuffer, C. U., & Kunde, W. (2014). Perceiving by proxy: Effect-based action control with unperceivable effects. *Cognition*, *132*(3), 251–261. <https://doi.org/10.1016/j.cognition.2014.04.012>
- Pfister, R., Weller, L., Dignath, D., & Kunde, W. (2017). What or when? The impact of anticipated social action effects is driven by action-effect compatibility, not delay. *Attention, Perception & Psychophysics*, *79*(7), 2132–2142. <https://doi.org/10.3758/s13414-017-1371-0>
- Piaget, J. (1952). *The origins of intelligence in children*. New York, NY: International Universities Press.
- Pieczkolan, A., & Huestegge, L. (2017). Cross-modal action complexity: Action- and rule-related memory retrieval in dual-response control. *Frontiers in Psychology*, *8*, 529. <https://doi.org/10.3389/fpsyg.2017.00529>
- Pieczkolan, A., & Huestegge, L. (2014). Oculomotor dominance in multitasking: Mechanisms of conflict resolution in cross-modal action. *Journal of Vision*, *14*(13), 18. <https://doi.org/10.1167/14.13.18>
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. <https://doi.org/10.1080/00335558008248231>
- Prinz, W. (1987). Ideo-motor action. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on perception and action* (pp. 47–76). Hillsdale, NJ: Erlbaum.
- Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, *9*(2), 129–154. <https://doi.org/10.1080/713752551>
- Proctor, R. W., & Reeve, T. G. (1990). Research on stimulus-response compatibility: Toward a comprehensive account. In R. W. Proctor & T. G. Reeve (Eds.), *Advances in Psychology. Stimulus-response compatibility* (Vol. 65, pp. 483–494). Amsterdam, Netherlands: North Holland. [https://doi.org/10.1016/S0166-4115\(08\)61234-0](https://doi.org/10.1016/S0166-4115(08)61234-0)
- Rantala, J., Majaranta, P., Kangas, J., Isokoski, P., Akkil, D., Špakov, O., & Raisamo, R. (2020). Gaze interaction with vibrotactile feedback: Review and design guidelines. *Human-Computer Interaction*, *35*(1), 1–39. <https://doi.org/10.1080/07370024.2017.1306444>
- Redcay, E., Dodell-Feder, D., Pearrow, M. J., Mavros, P. L., Kleiner, M., Gabrieli, J. D. E., & Saxe, R. (2010). Live face-to-face interaction during fMRI: A new tool for social cognitive neuroscience. *NeuroImage*, *50*(4), 1639–1647. <https://doi.org/10.1016/j.neuroimage.2010.01.052>
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York, NY: Appleton-Century-Crofts.
- Rescorla, R. A. (1968). Probability of shock in the presence and absence of cs in fear conditioning. *Journal of Comparative and Physiological Psychology*, *66*(1), 1–5. <https://doi.org/10.1037/h0025984>
- Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist*, *43*(3), 151–160. <https://doi.org/10.1037/0003-066X.43.3.151>
- Ricciardelli, P., Baylis, G., & Driver, J. (2000). The positive and negative of human expertise in gaze perception. *Cognition*, *77*(1), B1–B14. [https://doi.org/10.1016/S0010-0277\(00\)00092-5](https://doi.org/10.1016/S0010-0277(00)00092-5)
- Riechelmann, E., Weller, L., Huestegge, L., Böckler, A., & Pfister, R. (2019). Revisiting intersubjective action-effect binding: No evidence for social moderators. *Attention, Perception & Psychophysics*, *81*(6), 1991–2002. <https://doi.org/10.3758/s13414-019-01715-6>
- Rieger, M. (2007). Letters as visual action-effects in skilled typing. *Acta Psychologica*, *126*(2), 138–153. <https://doi.org/10.1016/j.actpsy.2006.11.006>
- Ristic, J., Friesen, C. K., & Kingstone, A. (2002). Are eyes special?: It depends on how you look at it. *Psychonomic Bulletin & Review*, *9*(3), 507–513. <https://doi.org/10.3758/BF03196306>
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169–192. <https://doi.org/10.1146/annurev.neuro.27.070203.144230>
- Rochat, P. (1998). Self-perception and action in infancy. *Experimental Brain Research*, *123*(1-2), 102–109. <https://doi.org/10.1007/s002210050550>
- Rolfs, M., Jonikaitis, D., Deubel, H., & Cavanagh, P. (2011). Predictive remapping of attention across eye movements. *Nature Neuroscience*, *14*(2), 252–256. <https://doi.org/10.1038/nn.2711>

- Rosenbaum, D. A. (2005). The Cinderella of psychology: The neglect of motor control in the science of mental life and behavior. *The American Psychologist*, *60*(4), 308–317. <https://doi.org/10.1037/0003-066X.60.4.308>
- Rothkirch, M., Ostendorf, F., Sax, A.-L., & Sterzer, P. (2013). The influence of motivational salience on saccade latencies. *Experimental Brain Research*, *224*(1), 35–47. <https://doi.org/10.1007/s00221-012-3284-4>
- Rubo, M., Huestegge, L., & Gamer, M. (2020). Social anxiety modulates visual exploration in real life - but not in the laboratory. *British Journal of Psychology*, *111*(2), 233–245. <https://doi.org/10.1111/bjop.12396>
- Saslow, M. G. (1967). Latency for saccadic eye movement. *Journal of the Optical Society of America*, *57*(8), 1030. <https://doi.org/10.1364/JOSA.57.001030>
- Sato, A., & Itakura, S. (2013). Intersubjective action-effect binding: Eye contact modulates acquisition of bidirectional association between our and others' actions. *Cognition*, *127*(3), 383–390. <https://doi.org/10.1016/j.cognition.2013.02.010>
- Schilbach, L. (2014). On the relationship of online and offline social cognition. *Frontiers in Human Neuroscience*, *8*, 278. <https://doi.org/10.3389/fnhum.2014.00278>
- Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *The Behavioral and Brain Sciences*, *36*(4), 393–414. <https://doi.org/10.1017/S0140525X12000660>
- Schulze, L., Renneberg, B., & Lobmaier, J. S. (2013). Gaze perception in social anxiety and social anxiety disorder. *Frontiers in Human Neuroscience*, *7*, 872. <https://doi.org/10.3389/fnhum.2013.00872>
- Schütz, A. C., Trommershäuser, J., & Gegenfurtner, K. R. (2012). Dynamic integration of information about salience and value for saccadic eye movements. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(19), 7547–7552. <https://doi.org/10.1073/pnas.1115638109>
- Senju, A., Hasegawa, T., & Tojo, Y. (2005). Does perceived direct gaze boost detection in adults and children with and without autism? The stare-in-the-crowd effect revisited. *Visual Cognition*, *12*(8), 1474–1496. <https://doi.org/10.1080/13506280444000797>
- Senju, A., & Johnson, M. H. (2009). The eye contact effect: Mechanisms and development. *Trends in Cognitive Sciences*, *13*(3), 127–134. <https://doi.org/10.1016/j.tics.2008.11.009>
- Shiffrar, M., & Freyd, J. J. (1990). Apparent motion of the human body. *Psychological Science*, *1*(4), 257–264. <https://doi.org/10.1111/j.1467-9280.1990.tb00210.x>
- Shin, Y. K., Proctor, R. W., & Capaldi, E. J. (2010). A review of contemporary ideomotor theory. *Psychological Bulletin*, *136*(6), 943–974. <https://doi.org/10.1037/a0020541>
- Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing. In R. W. Proctor & T. G. Reeve (Eds.), *Advances in Psychology. Stimulus-response compatibility* (Vol. 65, pp. 31–86). Amsterdam, Netherlands: North Holland. [https://doi.org/10.1016/S0166-4115\(08\)61218-2](https://doi.org/10.1016/S0166-4115(08)61218-2)
- Slobodenyuk, N. (2016). Towards cognitively grounded gaze-controlled interfaces. *Personal and Ubiquitous Computing*, *20*(6), 1035–1047. <https://doi.org/10.1007/s00779-016-0970-4>
- Sprague, N., & Ballard, D. (2003). Eye movements for reward maximization. In S. Thrun, L. K. Saul, & B. Scholkopf (Eds.), *Advances in neural information processing systems* (pp. 1467–1474). Cambridge, MA: MIT Press.
- Stephenson, L. J., Edwards, S. G., Howard, E. E., & Bayliss, A. P. (2018). Eyes that bind us: Gaze leading induces an implicit sense of agency. *Cognition*, *172*, 124–133. <https://doi.org/10.1016/j.cognition.2017.12.011>
- Tanriverdi, V., & Jacob, R. J. K. (2000). Interacting with eye movements in virtual environments. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 265–272. <https://doi.org/10.1145/332040.332443>
- Theeuwes, J., & Belopolsky, A. V. (2012). Reward grabs the eye: Oculomotor capture by rewarding stimuli. *Vision Research*, *74*, 80–85. <https://doi.org/10.1016/j.visres.2012.07.024>

- Tipples, J. (2002). Eye gaze is not unique: Automatic orienting in response to uninformative arrows. *Psychonomic Bulletin & Review*, *9*(2), 314–318.
- Tipples, J. (2005). Orienting to eye gaze and face processing. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(5), 843–856. <https://doi.org/10.1037/0096-1523.31.5.843>
- Tipples, J. (2008). Orienting to counterpredictive gaze and arrow cues. *Perception & Psychophysics*, *70*(1), 77–87. <https://doi.org/10.3758/pp.70.1.77>
- Tlauka, M., & McKenna, F. P. (1998). Mental imagery yields stimulus-response compatibility. *Acta Psychologica*, *98*(1), 67–79. [https://doi.org/10.1016/S0001-6918\(97\)00050-4](https://doi.org/10.1016/S0001-6918(97)00050-4)
- Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. Dunham (Eds.), *Joint attention: Its origins and role in development*. Hillsdale, NJ: Erlbaum.
- Urbina, M. H., & Huckauf, A. (2008). On object selection in gaze controlled environments. *Journal of Eye Movement Research*, *2*(4), 1–7. <https://doi.org/10.16910/JEMR.2.4.4>
- Van der Kamp, J., & Sundstedt, V. (2011). Gaze and voice controlled drawing. *Proceedings of the 1st Conference on Novel Gaze-Controlled Application*, 1–8. <https://doi.org/10.1145/1983302.1983311>
- Van der Wel, R. P., Welsh, T., & Böckler, A. (2018). Talking heads or talking eyes? Effects of head orientation and sudden onset gaze cues on attention capture. *Attention, Perception & Psychophysics*, *80*(1), 1–6. <https://doi.org/10.3758/s13414-017-1462-y>
- Vecera, S. P., & Johnson, M. H. (1995). Gaze detection and the cortical processing of faces: Evidence from infants and adults. *Visual Cognition*, *2*(1), 59–87. <https://doi.org/10.1080/13506289508401722>
- Venetti, A., Smith, T. J., & Senju, A. (2017). Gaze-contingent reinforcement learning reveals incentive value of social signals in young children and adults. *Proceedings of the Royal Society B*, *284*(1850), 20162747. <https://doi.org/10.1098/rspb.2016.2747>
- Verschoor, S. A., Paulus, M., Spape, M., Biro, S., & Hommel, B. (2015). The developing cognitive substrate of sequential action control in 9- to 12-month-olds: Evidence for concurrent activation models. *Cognition*, *138*, 64–78. <https://doi.org/10.1016/j.cognition.2015.01.005>
- Verschoor, S. A., Spape, M., Biro, S., & Hommel, B. (2013). From outcome prediction to action selection: Developmental change in the role of action-effect bindings. *Developmental Science*, *16*(6), 801–814. <https://doi.org/10.1111/desc.12085>
- Vogeley, K., & Roepstorff, A. (2009). Contextualising culture and social cognition. *Trends in Cognitive Sciences*, *13*(12), 511–516. <https://doi.org/10.1016/j.tics.2009.09.006>
- Vu, K.-P. L., & Proctor, R. W. (2004). Mixing compatible and incompatible mappings: Elimination, reduction, and enhancement of spatial compatibility effects. *The Quarterly Journal of Experimental Psychology*, *57*(3), 539–556. <https://doi.org/10.1080/02724980343000387>
- Walker, R., Husain, M., Hodgson, T.L., Harrison, J., & Kennard, C. (1998). Saccadic eye movement and working memory deficits following damage to human prefrontal cortex. *Neuropsychologia*, *36*(11), 1141–1159. [https://doi.org/10.1016/S0028-3932\(98\)00004-9](https://doi.org/10.1016/S0028-3932(98)00004-9)
- Wang, Q., Bolhuis, J., Rothkopf, C. A., Kolling, T., Knopf, M., & Triesch, J. (2012). Infants in control: Rapid anticipation of action outcomes in a gaze-contingent paradigm. *PLoS One*, *7*(2), e30884. <https://doi.org/10.1371/journal.pone.0030884>
- Wass, S., Porayska-Pomsta, K., & Johnson, M. H. (2011). Training attentional control in infancy. *Current Biology*, *21*(18), 1543–1547. <https://doi.org/10.1016/j.cub.2011.08.004>
- Watanabe, K., Lauwereyns, J., & Hikosaka, O. (2003). Effects of motivational conflicts on visually elicited saccades in monkeys. *Experimental Brain Research*, *152*(3), 361–367. <https://doi.org/10.1007/s00221-003-1555-9>
- Weeks, J. W., Howell, A. N., & Goldin, P. R. (2013). Gaze avoidance in social anxiety disorder. *Depression and Anxiety*, *30*(8), 749–756. <https://doi.org/10.1002/da.22146>
- Weiss, K., Schneider, W. X., & Herwig, A. (2014). Associating peripheral and foveal visual input across saccades: A default mode of the human visual system? *Journal of Vision*, *14*(11). <https://doi.org/10.1167/14.11.7>
- Weller, L., Kunde, W., & Pfister, R. (2017). Non-action effect binding: A critical re-assessment. *Acta Psychologica*, *180*, 137–146. <https://doi.org/10.1016/j.actpsy.2017.09.001>

- Weller, L., Pfister, R., & Kunde, W. (2019). Sociomotor actions: Anticipated partner responses are primarily represented in terms of spatial, not anatomical features. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(8), 1104–1118. <https://doi.org/10.1037/xhp0000658>
- Weller, L., Schwarz, K. A., Kunde, W., & Pfister, R. (2017). Was it me? - Filling the interval between action and effects increases agency but not sensory attenuation. *Biological Psychology*, *123*, 241–249. <https://doi.org/10.1016/j.biopsycho.2016.12.015>
- Wenke, D., Waszak, F., & Haggard, P. (2009). Action selection and action awareness. *Psychological Research*, *73*(4), 602–612. <https://doi.org/10.1007/s00426-009-0240-4>
- Wirth, R., Janczyk, M., & Kunde, W. (2018). Effect monitoring in dual-task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(4), 553–571. <https://doi.org/10.1037/xlm0000474>
- Wirth, R., Pfister, R., Brandes, J., & Kunde, W. (2016). Stroking me softly: Body-related effects in effect-based action control. *Attention, Perception & Psychophysics*, *78*(6), 1755–1770. <https://doi.org/10.3758/s13414-016-1151-2>
- Wolf, T., Sebanz, N., & Knoblich, G. (2018). Joint action coordination in expert-novice pairs: Can experts predict novices' suboptimal timing? *Cognition*, *178*, 103–108. <https://doi.org/10.1016/j.cognition.2018.05.012>
- Wolfe, J. M., O'Neill, P., & Bennett, S. C. (1998). Why are there eccentricity effects in visual search? Visual and attentional hypotheses. *Perception & Psychophysics*, *60*(1), 140–156. <https://doi.org/10.3758/bf03211924>
- Wolfensteller, U., & Ruge, H. (2011). On the timescale of stimulus-based action-effect learning. *Quarterly Journal of Experimental Psychology*, *64*(7), 1273–1289. <https://doi.org/10.1080/17470218.2010.546417>
- Wühr, P., & Huestegge, L. (2010). The impact of social presence on voluntary and involuntary control of spatial attention. *Social Cognition*, *28*(2), 145–160. <https://doi.org/10.1521/soco.2010.28.2.145>
- Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(5), 601–621. <https://doi.org/10.1037/0096-1523.10.5.601>
- Yarbus, A. L. (1967). *Eye Movements and Vision*. Boston, MA: Springer US. <https://doi.org/10.1007/978-1-4899-5379-7>
- Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology: General*, *81*(1), 141–145. <https://doi.org/10.1037/h0027474>
- Zaki, J., Kallman, S., Wimmer, G. E., Ochsner, K., & Shohamy, D. (2016). Social cognition as reinforcement learning: Feedback modulates emotion inference. *Journal of Cognitive Neuroscience*, *28*(9), 1270–1282. [https://doi.org/10.1162/jocn\\_a\\_00978](https://doi.org/10.1162/jocn_a_00978)
- Zaki, J., & Ochsner, K. (2009). The need for a cognitive neuroscience of naturalistic social cognition. *Annals of the New York Academy of Sciences*, *1167*, 16–30. <https://doi.org/10.1111/j.1749-6632.2009.04601.x>
- Ziessler, M., & Nattkemper, D. (2011). The temporal dynamics of effect anticipation in course of action planning. *Quarterly Journal of Experimental Psychology*, *64*(7), 1305–1326. <https://doi.org/10.1080/17470218.2011.553067>



# APPENDIX

## A ACKNOWLEDGEMENTS

Während ich diese Zeilen der Danksagung verfasse, befindet sich Deutschland in einem Ausnahmezustand. Es gelten Ausgangsbeschränkungen, die das Grundrecht auf Freiheit der Person mit dem Ziel der Bekämpfung einer Pandemie einschränken. Ein Zustand, der für meine Generation, die in ein vereintes und freiheitliches Deutschland geboren wurde, bis vor kurzem noch unvorstellbar war. In solchen Zeiten wird deutlich, wie wertvoll es ist, in einem Land zu leben, in welchem Freiheit nicht nur in Form von Freiheit der Person, sondern auch in Form von Freiheit der Wissenschaft, Forschung und Lehre als Grundrecht geschützt ist – eine Form der Freiheit, die kein Virus einzuschränken schafft. Im Gegenteil, gerade jetzt ist es für die Gesellschaft wichtig, auf ein effektives und funktionales Wissenschaftssystem zurückgreifen zu können, sodass schnellstmöglich ein Wirk- oder Impfstoff entwickelt werden kann, um die Ausbreitung des Virus einzudämmen. Daher ist es gerade in solchen Tagen sehr passend, sich die Bedeutung davon bewusst zu machen, in einem Land zu leben, in welchem Freiheit einen zentralen Wert darstellt.

An dieser Stelle möchte ich meinen Dank aussprechen an alle, die mich auf dem Weg hin zu und durch diese Arbeit begleitet haben. Die hier vorgestellten Ergebnisse wären so nicht entstanden, wenn ich nicht von einer Vielzahl von Menschen auf unterschiedlichste Art und Weise unterstützt worden wäre.

Als Erstes möchte ich mich bei meinem Erstbetreuer Prof. Dr. Lynn Huestegge bedanken, der in mir als Masterstudentin die Neugier geweckt hat, mich mit dem spannenden Feld der visuellen Kognition zu beschäftigen, und von Beginn an großes Vertrauen in mich setzte. Mit einer hervorragenden Kombination aus wissenschaftlicher Expertise und Ermutigung zu Selbstmanagement stand

er mir bei Fragen und Problemen jeder Art zur Verfügung. Prof. Dr. Huestegge besitzt die Fähigkeit, dass man nach einem Gespräch mit ihm nicht mehr problem-, sondern lösungs- und wachstumsorientiert denkt, sodass alles, was einem zuvor unmöglich zu schaffen erschien, plötzlich erreichbar wirkt. Danke!

Daneben gilt mein Dank meinem Zweitbetreuer Prof. Dr. Wilfried Kunde, der mit seiner Expertise, vor allem im Bereich der Ideomotorik, immer für fruchtbare Diskussionen und wertvolle Literaturtipps zur Verfügung stand. Des Weiteren danke ich Prof. Dr. Matthias Gamer und Prof. Dr. Erhard Wischmeyer für die Bereitschaft, mich als zusätzliche Mitglieder meines Promotionskomitees der GSLS (Graduate School of Life Sciences) durch die Zeit der Promotion zu begleiten.

Ich danke der Deutschen Forschungsgemeinschaft (DFG), die es mit ihren finanziellen Mitteln erst ermöglicht hat, diese Arbeiten durchzuführen sowie die Ergebnisse auf (inter-)nationalen Kongressen vorzustellen.

Was wäre man ohne die Unterstützung seiner Kollegen? An dieser Stelle möchte ich mich bei meinen (ehemaligen) Teamkollegen vom Lehrstuhl für Methoden, Kognition und Anwendung sowie bei meinen Mitautoren bedanken, auf die ich mich immer verlassen konnte und die mir mit Rat und Tat zur Verfügung standen. Ihr alle habt mit eurer Unterstützung zum Gelingen meiner Arbeit und zur Aufrechterhaltung einer angenehmen Arbeitsatmosphäre beigetragen, an die ich mich gerne zurückerinnere. Danke an PD Dr. Arvid Herwig, Prof. Dr. Gernot Horstmann, Jun.-Prof. Dr. Anne Böckler-Raettig, Dr. Aleksandra Pieczykolan, Dr. Mareike Hoffmann, Dr. Michael Steinborn, Dr. Tim Raettig, Dr. Lisa Weller, Dr. Roxana Kotter und Christina Breil. Mein besonderer Dank gilt Dr. Rainer Scheuchenpflug, der mit seiner humorvollen Art bereits in der Statistikvorlesung im ersten Semester meine Begeisterung für kritisches Denken und Hinterfragen weckte. In meinen Dank schließe ich auch alle studentischen Hilfskräfte am Lehrstuhl für Methoden, Kognition und Anwendung sowie meine Bachelorstudentin Eva Landmann ein, die mir durch ihren Fleiß und ihre Zuverlässigkeit eine große Unterstützung waren.

Mein Dank gilt allen weiteren aktuellen und ehemaligen Mitgliedern des Lehrstuhls für Psychologie III, für wissenschaftlichen Rat und Unterstützung, für ermutigende Worte, für sportliche

Mittagspausen, für gemeinschaftliche Spaghetti-Mittagspausen, für stärkende Kaffeepausen und natürlich für bereichernde Konferenzerfahrungen. Einen besonderen Dank möchte ich an dieser Stelle an Dr. Lisa Weller und Dr. Roland Pfister aussprechen, mit denen mich die Begeisterung für die Soziomotorik verbindet.

Ich danke dem Team des Sekretariats, Frau Barbara Galka, Frau Silvia Jahn und Frau Andrea Albert, für ihr offenes Ohr und ihre Unterstützung in jeglichen organisatorischen Fragen. Vielen Dank auch an Frau Katrin Lichosik und das Team der GSLS, die zusammen einen reibungslosen Ablauf der organisatorischen Prozesse auf meinem Weg zur Promotion ermöglichten.

Besonderer Dank gilt Anna Wagner, Markus Wagner, Dr. Nikola Kaul und Andreas Kaul für ihre Bereitschaft, mir als Fotomodelle für die in meiner Arbeit verwendeten Gesichts-Stimuli zur Verfügung zu stehen.

An dieser Stelle möchte ich gerne meinen Eltern Karin und Adolf Riechelmann dafür danken, dass sie mich von klein auf ermutigt und es mir ermöglicht haben, neugierig und wissbegierig durchs Leben zu gehen, meine Interessen zu verfolgen und Herausforderungen anzunehmen. Mein Dank gilt auch meinen Geschwistern Anna Wagner und Andreas Kaul sowie deren Familien. Als große Geschwister wart ihr mir Vorbild und Ansporn zugleich und habt mich dazu ermutigt, mit Ehrgeiz meine Ziele zu verfolgen.

Mein ganz besonderer Dank aber gilt Dr. Andreas Steinbacher, ohne dessen liebevolles Verständnis und unerschöpfliche Geduld diese Arbeit niemals hätte gelingen können. Von Anfang an hat er an mich geglaubt und mir den notwendigen Halt und Beistand gegeben, wenn ich selbst ins Zweifeln geriet. Sein kritischer Geist, seine Diskussionsfreude und seine differenzierten Anmerkungen bei der Durchsicht dieser Arbeit haben einen wichtigen Beitrag zur Anfertigung und Vollendung meiner Dissertation geleistet. Ihm widme ich diese Arbeit voller Vorfreude auf den gemeinsamen Weg, der vor uns liegt.





## C AFFIDAVIT (EIDESSTATTLICHE ERKLÄRUNG)

I hereby confirm that my thesis entitled

“GAZE INTERACTION: COGNITIVE MECHANISMS OF OCULOMOTOR ACTION CONTROL”

is the result of my own work. I did not receive any help or support from commercial consultants. All sources and/or materials applied are listed and specified in the thesis.

Furthermore, I confirm that this thesis has not yet been submitted as part of another examination process neither in identical nor in similar form.

Würzburg, May 8, 2020

---

Eva Katharina Riechelmann

Hiermit erkläre ich an Eides statt, die Dissertation

„BLICKINTERAKTION: KOGNITIVE MECHANISMEN DER OKULOMOTORISCHEN  
HANDLUNGSKONTROLLE“

eigenständig, d.h. insbesondere selbständig und ohne Hilfe eines kommerziellen Promotionsberaters, angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

Ich erkläre außerdem, dass die Dissertation weder in gleicher noch in ähnlicher Form bereits in einem anderen Prüfungsverfahren vorgelegen hat.

Würzburg, 8. Mai 2020

---

Eva Katharina Riechelmann