The Effect of Embodiment, Mindfulness, and Stereotype Threat on Mental Rotation Performance and Cognitive Load

Inaugural-Dissertation zur Erlangung der Doktorwürde der

Fakultät für Humanwissenschaften der Universität Regensburg,

vorgelegt von

Robert Bauer
aus Augsburg
-2023-

Regensburg, 2023

Gutachterin (Betreuerin): Prof. Dr. Petra Jansen

Gutachter: Prof. Dr. Dr. h.c. Hans Gruber

Content

List of Tables

List of Figures

Abbreviations and Acronyms

This list provides information regarding important and reoccurring abbreviations and acronyms:

ACC	Accuracy; rate of accuracy
ASS	"Anstrengungsskala Sport"; Likert-type scale to rate perceived exertion
BP	Body postures; part of the variable "stimulus type"
C	Control group; control condition
CF	Cube figures; part of the variable "stimulus type"
DEG	Angular disparity between two figures shown on a screen; main moderator of difficulty in mental rotation tasks
F	Female; women
HF	Human figures; part of the variable "stimulus type"
M	Male; men
MF	Mindfulness; mindfulness group; mindfulness condition
PD	Difference between the maximum pupil diameter and the respective baseline pupil diameter for each trial as a measure of cognitive load
PP	Pre- and posttest experimental setup
RPE	Rate of perceived exertion
RT	Reaction time
RT-Sum	Total sum of reaction time for all items for each block (pre- and posttest)
SEX	Variable and effect of sex and/or gender differences. Within the scope of many experiments, it is impossible to distinguish between sex and gender in the analyses (see **Section 7.3** and **7.4.3**). Hence, both terms are often used interchangeably in existing literature (see **Section 2.5**), which also applies to this thesis.
STEM	Fields of science, technology, engineering, and math
STI	Stimulus type; variable refers to cube figures, human figures, and body postures
TMS	Toronto Mindfulness Scale

Acknowledgements

The Road goes ever on and on
Down from the door where it began.
Now far ahead the Road has gone,
And I must follow, if I can,
Pursuing it with eager feet,
Until it joins some larger way
Where many paths and errands meet.
And whither then? I cannot say.

J. R. R. Tolkien (2005)

I am very content reaching the end of this road that is my Ph.D. and I am looking forward to all the possible paths ahead of me. Like on every good journey, many valued people accompanied and supported me. While trying to keep it short, I would like to express my appreciation to you all.

First of all, I am very grateful to my supervisor Prof. Dr. Petra Jansen for all the great support. Petra, I really appreciate that you gave me the freedom to choose my own topic while providing sound advice regarding possible options. Thank you for your professional expertise and support regarding the process of developing an experimental set-up, analyzing the data, and writing this thesis. Your kind and patient guidance made it very delightful to work with you. Your great work attitude and passion for science will continue to inspire me.

Thanks to the whole team of the Institute of Sports Science, that is, scientists and all other personnel. Everyone added a unique and friendly personality that combined to an amicable and respectful work environment. With such a great team, going to work was always very enjoyable.

My deepest gratitude and appreciation go to my dear parents Beate and Werner. Danke, dass ihr es mir ermöglicht habt zu studieren, ins Ausland zu gehen und meinen Weg zu meistern. Eure immerwährende positive Unterstützung und bedingungslose Liebe sind fest in meinem Herzen verankert.

I also thank my beloved sister Sanny, my brother-in-law Abraham, and their little rascals Tom and Oli. Ich danke dir, Sanny, dass du solch ein positiver und herzenswarmer Mensch bist. Als Schwester hast du mir gezeigt, mit einem Lachen und Sonne im Herzen durch das Leben zu gehen. Gracias, Abraham, por ser un genial cuñado y amigo a la vez. No podría imaginarme alguien mejor para mi hermana. Danke an die zwei Knödel, welche unsere fröhliche Familie noch glücklicher gemacht haben. Ich freue mich darauf, viel mit euch zu unternehmen.

I deeply thank my girlfriend Marta. Gracias por tu apoyo, tu risa y tu amor. Te adoro, porque contigo el tiempo se convierte en pura vida. Tú eres la respuesta a mi miau. También quiero decir gracias a Isa y Javi por ser los mejores suegritos imaginables.

My thanks go to my friends Claudio and Florian, who read the final manuscript and gave me helpful feedback from a mathematician’s and physicist’s viewpoint.

I want to thank my friend Laura. Gracias por tus consejos y tu apoyo en el mundo científico. Te cogí cariño desde el primer momento y me alegro de tenerte como amiga.

Lastly, thanks to everyone who has shaped me directly or indirectly into the person I am today. I’ll be seeing you on the roads to come.

Summary

Spatial abilities play an important role in everyday life, as they are essential for understanding and interacting with our surroundings like navigation or handling tools (Newcombe & Shipley, 2015). They are also linked to success in STEM fields, such as mathematics (Xie et al., 2020). One of the spatial abilities is mental rotation, which can be described as the cognitive ability to visualize and compare rotated objects in the mind. Mental rotation is an important skill and a necessity in various high-paying occupations, such as surgery, architecture, or engineering. Due to a persisting gender gap in STEM, the underlying reasons for sex differences in mental rotation performance continue to be an important research subject. With regard to the gender gap, possible influencing factors like gender stereotypes and resulting stereotype threats are investigated. Furthermore, methods to ameliorate their negative effects are progressively developed and researched. A promising method derives from embodied cognition. Embodiment effects have been shown to improve cognitive performance (e.g., Rahe & Jansen, 2023) and to counter the negative influence of stereotype threats (Weger et al., 2012). They can be elicited via integrated aspects of a cognitive test itself (e.g., stimuli; Campbell et al., 2018) or via separate interventions (e.g., mindfulness; Rahe & Jansen, 2023). These possibilities are explored in this thesis.

Over the course of three studies, we investigated the effects of embodiment—via embodied stimuli and embodied mindfulness—and stereotype threat on chronometric mental rotation task performance and cognitive load. Possible sex differences were also included. In the first study, we analyzed mental rotation task performance with three types of stimuli that had different levels of embodiment, and gender. Here, pupillometric measurements were taken with eye-tracking technology to analyze cognitive load. In the second study, the effects of a 20-minute embodied focused attention mindfulness meditation were analyzed. The meditation and a control treatment, and gender were investigated regarding mental rotation performance and subjective cognitive effort. In the third study, we used a shorter embodied focused attention mindfulness meditation than in the second study, and combined it with explicit stereotype threat applied through auditory induction. All conditions and gender were investigated regarding mental rotation performance and pupillometric measurements of cognitive load.

In total, the studies provided evidence that the effects elicited by embodied test stimuli as well as one-time mindfulness interventions affected men and women similarly in their behavioral performance and cognitive load. Regarding these variables, the thesis's findings also indicate that one brief mindfulness intervention may not be enough to induce significant embodiment effects. With the addition of a stereotype threat induction, the results of the third study indicate that, generally, higher levels of state mindfulness could prevent stereotype threat to influence cognitive performance. Lastly, the first study shows that the mental rotation of embodied figures is easier to process and needs less cognitive effort to solve the task.

In conclusion, this thesis provides new insights about possibly beneficial aspects of embodied cognition and mindfulness on mental rotation ability and cognitive load, and addresses sex differences and the influence of stereotype threat. The results may improve and inspire future research, and the discussion may highlight the practical importance due to the connection between mental rotation, mathematical abilities, and employment in the STEM fields.

1 Preface

In cognition, spatial abilities are an important set of skills that involve the mental representation and transformation of visual information such as objects or shapes, and the relationships between them. Apart from being important in everyday life, they seem to have a close connection to general intelligence. Spatial ability performance is a strong predictor of success in academic and professional areas of Science, Technology, Engineering, and Mathematics disciplines. As one of the spatial abilities, mental rotation is a skill that is used very often, for instance, when interpreting a map or when trying to figure out how a piece of luggage would fit best into a fixed space in a car. In some forms of mental rotation tasks, there is a well-established sex difference in favor of men over women. As mental rotation competence is a necessity in various high-paying occupations, such as surgery, architecture, or engineering, the underlying reasons for possible sex differences in mental rotation performance continue to be an important research subject.

Mental rotation can be described as the cognitive ability to visualize and compare rotated objects in the mind. This involves the "ability to imagine how an object that has been seen from one perspective would look if it were rotated in space into a new orientation and viewed from the new perspective" (Johnson & Moore, 2020, p.1). Since the seminal study of R.N. Shepard and Metzler in 1971, mental rotation has been increasingly investigated in thousands of studies. This interest is also not subsiding, as mental rotation research has still many open questions to answer. These can range from general cognitive insights to practical applications, such as the intention to mitigate gender differences, which could eventually lead to gender equality and equity.

Therefore, the analysis of sex differences in spatial abilities is an important topic. Various possible factors have been identified and analyzed in mental rotation experiments. They can range from factors related to biological sex to social factors related to gender. One of the latter are gender stereotypes and corresponding stereotype threats. For that reason, the possibilities to circumvent the negative effects of such stereotypes are important to investigate and develop. One promising candidate could be different forms of meditation practices, such as focused attention meditation. Here, the mind directs and maintains focus on one specific object (e.g., breath) over a defined time period to decrease mind wandering and to increase attentional control. The effects of such meditation practices are closely connected to embodiment. Embodiment can be described as the idea that cognitive processes are grounded in the organism's sensory and motor experiences through bodily interactions with the world, which lead to direct effects of bodily experiences on the mind (Barsalou, 2008). As indicated, embodiment can have an influence through treatments, but it can also be encompassed by parts of a test itself, such as the test stimuli. Both aspects can be analyzed via a biological marker—the measurement of pupil size—that can give insight into cognitive changes evoked by embodiment and might provide new information regarding the genders. As demonstrated, sex and gender differences in mental rotation performance, embodiment and embodied mindfulness, and stereotype threat are connected. This thesis attempts to contribute to our understanding of the mental rotation ability by investigating these three aspects.

2 Theoretical Background and State of Research

The following section introduces spatial abilities in general and mental rotation more specifically. It will further introduce possible influence factors regarding mental rotation performance. Finally, the connection of mental rotation and cognitive effort will be presented.

2.1 Spatial Abilities

A specific class of cognitive abilities is spatial abilities. In everyday activities, spatial abilities play an important role like in tool use and navigation (Newcombe & Shipley, 2015). Research indicates that spatial abilities predict achievement and attainment in science, technology, engineering, and mathematics (STEM) disciplines (Humphreys et al., 1993; Shea et al., 2001; Wai et al., 2009; Xie et al., 2020). For instance, spatial abilities play an important role for adolescents' choice of advanced education in STEM subjects (Wai et al., 2009).

McGrew (2009) defined spatial ability as the ability to generate, store, retrieve, and transform visual information. Linn and Petersen (1985) tried to classify spatial abilities into three parts: spatial perception, spatial visualization, and a separate category for mental rotation. Other authors added more (sub-)categories such as spatial orientation or perceptual speed and proposed different titles for the same or similar categories (see Harle & Towns, 2011; Yilmaz, 2017). However, the definition of and distinction between these groups was not completely clear (Hegarty & Waller, 2005; Kozhevnikov & Hegarty, 2001; Xie et al., 2020). As a consequence, Newcombe and Shipley (2015), and Uttal et al. (2013) proposed to distinguish between fundamental task characteristics and differentiated spatial abilities according to two dimensions: intrinsic versus extrinsic and static versus dynamic. Intrinsic spatial information describes the spatial properties of an object and its parts, whereas extrinsic information describes an object regarding its location in relation to a frame or reference, and to other objects. The second dimension differentiates between static and dynamic tasks, that is, whether the object or objects are being moved and change their position or orientation, either in physical actuality or through mental simulation (Newcombe & Shipley, 2015; Uttal et al., 2013; Xie et al., 2020).

Within this framework, intrinsic-dynamic tasks require a mental transformation of the spatial coding of an object, such as a rotation in the Mental Rotation Test (R. N. Shepard & Metzler, 1971; Vandenberg & Kuse, 1978) or Cards Rotation Test (French et al., 1963). Other tests can include cross-sectioning, folding, and plastic deformations. In contrast, intrinsic-static tasks involve the identification of objects as members of categories, or the perception of objects amidst distracting information (Newcombe & Shipley, 2015). This applies, for example, in the embedded figures task, where one has to identify simple figures that are embedded in more complex figures (Witkin, 1971). In extrinsic-dynamic tasks, participants have to visualize a whole environment from a different position (Uttal et al., 2013). One example is Piaget's Three Mountains Task (Piaget & Inhedlerl, 1969). In this test, participants view a landscape on a table and have to choose the picture that most resembles this landscape, but from the point of view of the experimenter seated on the other side of it. Lastly, extrinsic-static tasks require the understanding of abstract spatial principles independent of the object in question such as the Rod and Frame Test, where a rod inside a tilted frame has to be considered with regard to its alignment with verticality (see Witkin & Asch, 1948; Zoccolotti et al., 1993).

This proposed 2x2 classification of Newcombe and Shipley (2015), and Uttal et al. (2013) was later questioned by Mix et al. (2018), in particular the dynamic-static dimension. In their confirmatory factor analysis, this dimension was not valid in three samples of spatial performance of kindergarten, third grade, and sixth grade students. In their review, Buckley et al. (2018) categorized dynamic spatial abilities into temporal processing, cognitive action, and moving stimuli. The latter defined that the stimulus within the spatial task is moving. Thus, tasks like mental rotation, paper folding, and block design should be considered to require intrinsic-static spatial abilities, opposed to their aforementioned intrinsic-dynamic categorization (Buckley et al., 2018; Xie et al., 2020). They further identified the additional category of visual-spatial memory, but also emphasized the necessity for further investigation of the distinction of static and dynamic spatial abilities.

Although a final definition and classification of spatial abilities has yet to be determined, they are already intensely studied for their practical implications, especially for being an important predictor of STEM performance (Buckley et al., 2018; Shea et al., 2001; Wai et al., 2009, 2010). For instance, they are related to specific abilities like problem solving (Geary et al., 2000) and mathematical ability (Xie et al., 2020). This relationship between mathematical and spatial ability is not linear, and applies for all categories defined by Uttal et al. (2013) and the additional category of visual-spatial memory (Xie et al., 2020). Uttal et al. (2013) pointed out that educational endeavors focus on reading and science, which both influence personal achievement. They also emphasized that the lack of training of basic skills like spatial thinking is a missed opportunity, as these skills are malleable and can underlie performance in various areas, especially STEM performance in particular (e.g., chemistry, Wu & Shah, 2004; or physics, Kozhevnikov et al., 2007), and in general (Gilbert, 2005), like reasoning about scientific diagrams (Stieff, 2007; Uttal et al., 2013). Therefore, while still missing definite classifications, spatial abilities are an important field in cognitive and developmental research.

2.2 Mental Rotation Ability

Mental rotation is a spatial ability, which is also strongly connected to performance, success, and education in STEM fields (e.g., Hausmann, 2014; Moè, 2016; Moè et al., 2018). For instance, it is required for mathematics (Hegarty & Kozhevnikov, 1999; Xie et al., 2020), problem solving (Geary et al., 2000), and science (Peters, Chisholm, et al., 1995). The term describes the cognitive ability to mentally rotate two- or three-dimensional objects or images fast and accurately (Linn & Petersen, 1985; Shepard & Metzler, 1971). In other words, it is the ability to imagine how an object would look like, if it were rotated from its original position. This means, spatial information can be represented mentally and then pivoted around all spatial axes (Shepard & Metzler, 1971).

Mental rotation is defined as an intrinsic-dynamic spatial ability, according to the classification of Uttal et al. (2013). Buckley et al. (2018) argued that it could also involve intrinsic-static spatial ability, as the stimuli are not moving. Furthermore, additional factors might be involved; for instance, within these tests, perspective taking seems to take place simultaneously with mental rotation strategies (Hegarty, 2018; Hegarty & Waller, 2004). Additionally, apart from the mental rotation process itself, mental rotation is a complex cognitive task that is processed in various stages. Heil and Rolke (2002) describe the perceptual stages (perceptual processing, identification and discrimination of stimuli, identification of orientation), stages of the rotation process itself (mental rotation, judgment of parity), and decision processing stages (response selection, execution) (see also R. N. Shepard & Cooper, 1986). To complete these stages, another necessary factor for mental rotation tasks is the visual working memory as storage of object representations (Hyun & Luck, 2007), which is addressed in more detail in **Section 2.9.3**.

2.3 Object-Based Versus Egocentric Transformations

Two mental rotation tasks are often used: an object-based or egocentric mental rotation. Both transformation types are connected to different cognitive processes. In object-based transformations, the observer's position remains fixed, and the object is rotated mentally in relation to the surrounding environment. For egocentric transformations, the observer is tasked to change their own perspective by an imaginary rotation of their own body (i.e., mentally rotate his own body relative to the object) (Voyer et al., 2017; Zacks et al., 2000).

Both transformations are normally represented by different tests. In an object-based transformation task, the participants have to decide whether two objects, which are rotated to each other, are mirror reversed to each other (then they are called "different") or not (then they are called "same"). In an egocentric transformation—for example, using human figures with raised arms (or hand pictures)—participants have to determine if it is the left or right arm (hand).

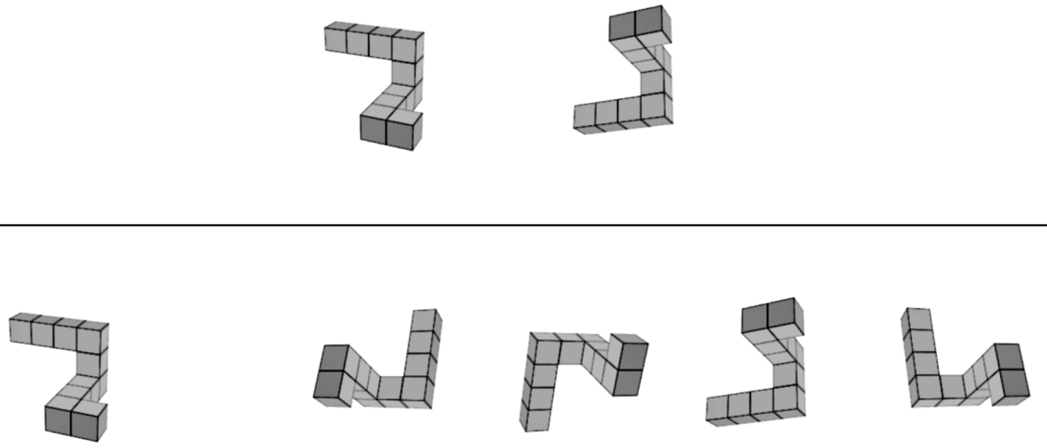
Unfortunately, egocentric and object-based mental rotation tasks confound the stimulus type (embodied versus non-embodied) and task instruction type (egocentric versus object-based). However, it has been demonstrated that not the type of stimuli but the kind of instruction (left/right vs. mirrored/non-mirrored) determined, if an egocentric or object-based transformation is evoked (Voyer et al., 2017). The two instruction types also reveal different test results (Voyer et al., 2017; Zacks et al., 2000). Because of these instruction differences, while being defined as egocentric mental rotation tests, egocentric tests are thus in separate categories in the classification of spatial abilities by Uttal et al. (2013) and Buckley et al. (2018) than object-based mental rotation tests. Therefore, in everything that follows, the main focus is on object-based mental rotation transformations.

2.4 Object-Based Mental Rotation Tests

To assess mental rotation performance, two types of mental rotation tests are commonly used. Chronometric tests (see Figure 1, top) are based on the seminal study of R. N. Shepard and Metzler (1971). Two objects are presented that are either the same (rotated) or different (mirrored). Psychometric tests (see Figure 1, bottom) are based on the study of Vandenberg and Kuse (1978). Here, one target object is compared to four alternatives. Two are the same (rotated) and the other two are different (mirrored or structurally different).

Figure 1

Examples of Trials of Chronometric (Top) and Psychometric (Bottom) Mental Rotation Tests.



Note. Abstract cube figures developed by Amorim et al. (2006); reprinted and adapted with permission by Michel-Ange Amorim.

2.4.1 Chronometric Tests

Chronometric mental rotation tests typically present two figures to the participants. These figures are either two-dimensional stimuli or three-dimensional stimuli rendered as two-dimensional images. Being instructed to be fast and accurate, participants have to determine, whether the figures are the same (i.e., one can be transformed into the other via a rotation) or different (i.e., mirrored, and neither can be transformed into the other via a rotation). For every item, the reaction time and the correctness of the answer are recorded, which can give insight into the cognitive process of each task. In their seminal study, Shepard and Metzler (1971) described that with abstract cube figures and object-based instructions, reaction times increase in a linear manner as a function of angular disparity. Mean reaction times increased from about one second with no rotation up to values between four and six seconds at 180° angular disparity for correctly answered non-mirrored pairs of cube figures. Later research has shown similar results with not always as perfect linear relationships. The relationship still illustrates that participants continuously rotate the presented objects in their mind with a more or less constant rotation velocity, which can be depicted by the linear slope of a time over angular disparity graph. At 0° angular disparity, the intercept of this graph depicts the necessary time to process two identical figures without rotation (Heil & Rolke, 2002). The relationship between error rates and angular disparity follows a weaker but somewhat similar pattern like reaction times with a constant and approximately linear increase of error rates with increasing rotation angles.

Although rotation speed seems to be directly linked to the rotation angle, the linear dependency of the accuracy rates cannot be easily explained similarly. Hyun and Luck (2007) proposed that during manipulations (even spatial transformations), visual form information is stored in the object memory system while being manipulated, instead of the spatial memory system. Hence, the mental representation of the rotation might interfere with storing the representation in the working memory (Jost, 2022). While being instructed to rotate as fast and accurately as possible, participants could still favor a speed-accuracy trade-off, that is, taking longer to solve more accurately. However, the effect illustrated by accuracy and speed-accuracy curves ensures that slower reaction times are not caused by speed-accuracy trade-offs (Carpenter et al., 1999; Hyun & Luck, 2007). To analyze the reaction times, chronometric mental rotation tests only include tasks with correct answers. This is usually done in cognitive tests measuring reaction times, as it is unknown why an error occurred in wrongly answered tasks. Consequently, the necessary amount of items presented to each participant increases to ensure that for each group, sufficient data are available, preserving statistical power.

For the stimulus items, three-dimensional cube figures are the most used ones. They were made popular by the seminal study of R. N. Shepard and Metzler (1971). The amount and organization of the connected cubes have been modified by scientists over the years. Peters and

Battista (2008) later created a comprehensive library of 28032 cube figures as a material pool for standardized stimuli, which has been widely used.

Since the early studies with chronometric mental rotation tests, the analysis of reaction time and accuracy data consists of a methodological problem. As mirrored figures cannot be brought into congruence, they are typically excluded from analysis (Jolicœur et al., 1985; R. N. Shepard & Metzler, 1971). As the tests usually consist of equally numbered mirrored and non-mirrored tests, this applies to half of the tasks. Their discarding means losing valuable research time as well as a noticeable amount of power (Brysbaert & Stevens, 2018). Additionally, the analysis of accuracy with only non-mirrored stimuli does not take into account false alarms and correct rejects, which might lead to different patterns than hits and misses (Jost & Jansen, 2020). By incorporating false alarms, it is possible to analyze sensitivity instead of accuracy (Stanislaw & Todorov, 1999). However, as false alarms would only occur with mirrored figures, they would not be congruent, and thus could not be matched to definite angles of rotation. That means they would only suffice for global analyzes of sensitivities (Jost & Jansen, 2020). To counteract the necessity to exclude half of the trials, Jost and Jansen (2020) proposed a new approach to analyze chronometric mental rotation tests. In their novel paradigm, they present three figures, with two being the base figures that are mirrored to each other, and one being the stimulus figure that is a rotated version of one of the former. Instead of the same/different decision making task from the typical tasks with two figures, participants are asked to find out, if the stimulus figure can be rotated into congruence with one or the other base figure. This only applies to one of these, and the rotation into congruence is therefore independent of stimulus orientation (Jost & Jansen, 2020). The study showed similar results like the Shepard and Metzler (1971) design, and was deemed applicable for measuring the mental rotation ability. The authors point out that it cannot yet be inferred by looking at the behavioral data alone whether similar strategies (e.g., actual mental rotation process) are applied like in the original design to solve the task. In accordance with Brysbaert and Stevens (2018), their results showed an increase of power for their design (Jost & Jansen, 2020).

2.4.2 Psychometric Tests

Vandenberg and Kuse (1978) developed the psychometric tests as paper and pencil tests by using similar cube figures as the chronometric tests. Their main advantage was the quick and simultaneous assessment of large subject groups. Here, a target object (normally placed on the left) has to be compared to four alternative objects. Two of these are rotated (but the same figure) and the other two are different (mirrored or structurally different figures). Test subjects have to select the two rotated figures in every trial. Peters, Laeng, et al. (1995) later created the most used versions to date, in which participants have to solve two sets of twelve tasks preceded by three practice trials. Each block of twelve tasks has to be completed within three minutes, while unanswered trials are considered as wrong answers. To reduce the impact of guessing items correctly, the test only credits one point per trial, when both correct alternatives are selected (Peters, Laeng, et al., 1995).

2.5 Sex Differences in Mental Rotation Test Performance

One important matter in mental rotation research is the potential existence of sex differences. The performance differences in psychometric mental rotation tests favoring men are one of the largest sex differences observed in cognitive psychology (Halpern, 1989, 2012; Voyer et al., 1995). Meta-analyses found sex differences in favor of boys and males throughout various developmental stages, that is, in childhood (Lauer et al., 2019), adolescence (Lauer et al., 2019; Voyer, 1995), and adulthood (Voyer et al., 1995). For that reason, a lot of research has been conducted to analyze the differences in spatial abilities and their connection to STEM performance, but definite causes are yet to be determined. In contrast to the more stable differences in psychometric tests, regarding the behavioral aspects, there is a continuous discussion about whether and how sex influences the performance in chronometric mental rotation tests. Here, differences are smaller and mostly non-significant, or emerge only in subtests or for specific stimuli (Jansen-Osmann & Heil, 2007; Peters & Battista, 2008; Rahe & Jansen, 2022). Overall, the various versions of mental rotation tests correlate in terms of performance, but their differing results regarding sex differences cannot be extensively explained, yet (Voyer et al., 2006).

However, if sex and/or gender differences appear, several explanation possibilities arise. These are biological factors, for instance, genetics (Quinn & Liben, 2014), sex hormone levels, menstrual cycle, sexual orientation (Alexander & Son, 2007; Grimshaw et al., 1995; Hausmann et al., 2000, 2009; Peters, Chisholm, et al., 1995; Peters et al., 2007), task performance mechanisms (e.g., rotational axis or time limit in tests; Gauthier et al., 2002; Peters, 2005; Voyer, 2011), or sociocultural factors like spatial toys choice and stimulus familiarity, socioeconomic factors, attending a single-sex school, or also gender stereotypes/stereotype threat (Baenninger & Newcombe, 1989; Hausmann et al., 2009; Heil et al., 2012; Moè & Pazzaglia, 2006; Rahe & Jansen, 2022, 2023; Ruthsatz et al., 2014, 2015, 2017; Steele & Aronson, 1995; Titze et al., 2011). Regarding the latter, it is also not clear, if performance differences stem from gender identification (social factor) or sex (biological factor). Research indicates differences related to both. Furthermore, the mental rotation performance of girls could be influenced by the preference for soccer over dancing (Pietsch et al., 2019), or exposure to mothers' spatial language use (Ralph et al., 2021; see Rahe & Jansen, 2022). Other possible factors are education and academic background (Moè et al., 2018; Peters, Chisholm, et al., 1995; Peters et al., 2007), and strategy selection (Hegarty, 2018; Heil & Jansen-Osmann, 2008; Nazareth et al., 2019; Voyer et al., 2020). Moè et al. (2018) found an effect of the use of spatial toys on mental rotation, with women of STEM degrees showing better performance, if they preferred spatial toys during their childhood. Similarly, Voyer et al. (2000) reported better mental rotation results for those of both genders, who had preferred spatial toys during their childhood. In the following sections regarding different influences on mental rotation performance, more information about possible sex differences within and between these will be provided.

2.6 Stereotype Threat in Spatial Tasks

Stereotype threat derives from gender stereotypes, which can be described as social phenomena that represent the beliefs about the characteristics of females and males (Rahe & Jansen, 2023). Stereotype threat can be defined as the “the risk of confirming, as self-characteristic, a negative stereotype about one’s group” (Steele & Aronson, 1995, p. 797). This can lead to a disruptive state (by triggering disruptive mental load) that undermines the performance in this specific domain, for example, in cognitive tests (Croizet et al., 2004; Spencer et al., 2016). In their meta-analysis, Doyle and Voyer (2016) found a significant effect of stereotype threat on women’s performance for math, but not for spatial tasks. They pointed out that stereotype manipulation might not be as effective when a stereotype is not widely accepted already. Aronson and McGlone (2009) also described that the experience of a stereotype threat can only start with the awareness that one is a member of a negatively stereotyped group. Stereotype threat can be measured implicitly by its effect on the dependent variables (e.g., Weger et al., 2012) or explicitly by using manipulation questionnaires with Likert-type scales (e.g., Neuburger et al., 2015; Pennington et al., 2019). As the counterpart to stereotype threat, stereotype lift means that for an individual, the own group is assumed to perform better in a certain topic (Walton & Cohen, 2003). For instance, females showed better results in spatial tasks that were manipulated with a stereotype lift (Doyle & Voyer, 2016).

Mental rotation performance can be influenced by these effects on several levels. Regarding gender differences in mental rotation and spatial abilities in general, stereotype threat and lift effects can be activated either directly by instructing the participants that one gender would outperform the other one (Heil et al., 2012), or indirectly by instructing participants to decide which gender would more likely outperform the other one in specific tasks (Hausmann et al., 2009), and by using gender-stereotyped material in the task (Rahe & Jansen, 2022; Rahe et al., 2021; Ruthsatz et al., 2017). In mental rotation, the stimuli themselves could often be regarded as more male stereotyped. Generally, spatial abilities are considered to be more male stereotyped, which might also be related to the STEM field interests (e.g., Law et al., 2021; Makarova et al., 2019). By using stereotyped objects in mental rotation tasks (e.g., dolls and cars), Ruthsatz et al. (2017) found a significant interaction between the gender and object type (female and male stereotyped) of children in the fourth grade, where male stimuli were more familiar to boys than female stimuli and vice versa. Investigating adults, Rahe et al. (2021) reported a significant interaction of gender and stereotyped material, with a significant gender differences for male stereotyped stimuli. In their other study, they used abstract cube figures and pellet figures (similar in shape, but replacing the cubes with round pellets) that were rotated in depth. For the cube figures, men showed better performance than women, but no interaction of gender and material was found (Rahe et al., 2021).

A lot of studies applied stereotype threat conditions via instructions—by telling participants that one gender outperforms the other—to measure their effects on mental rotation tasks performance and reported significant effects (e.g., Heil et al., 2012; Moè & Pazzaglia, 2006), and also investigated possibilities to reduce or prevent their influence (e.g., Hausmann, 2014; Martens et al., 2006). Moè and Pazzaglia (2006) showed that by inducing gender beliefs artificially, mental rotation task performance was affected. For both genders, the performance increased due to stressing the participants' *own* gender superiority and decreased by stressing the opposite gender superiority. By applying the same scoring method, these results were replicated by Heil et al. (2012). In their further analysis, they reported women performing according to gender belief induction, but their guessing behavior (willingness of guessing analyzed by comparing a liberal and conservative scoring method of a psychometric mental rotation test) was not affected. Men showed the opposite results, that is, their performance was unaffected, but the guessing behavior was in line with gender belief induction (Heil et al., 2012). In another study, Hausmann (2014) found female art students to be more susceptible to the influence of a threat condition than female science students. They concluded that being confident in one's own abilities is a better predictor of mental rotation performance than how one perceives a stereotype about the own gender. Furthermore, Martens et al. (2006) showed that by raising the awareness of females toward their positive characteristics and values would result in a lower impact of the negative effects through the stereotype threat. Lastly, stereotype threat activation only produced negative effects in women with feminine gender role identification (Tempel & Neumann, 2016). Rahe and Jansen (2023) summarized that having stronger beliefs about one's ability reduces the effects of an applied stereotype threat, and that the intended effects by this threat may generally not be consistent. They also pointed out that the mentioned beliefs could be improved by training of spatial activities. In several studies, spatial and motivational training improved the mental rotation task performance in women (e.g., Meneghetti et al., 2016; Moè, 2021; Neubauer et al., 2010; Sorby, 2009; see Rahe & Jansen, 2023). One possibility to change the beliefs in the own ability and a way to mitigate the effects of stereotype threat is the practice of mindfulness. In the upcoming sections, mindfulness will be introduced, and the effects of brief one-time interventions will be a major focus in **Study 2** and **Study 3** (here also in interaction with stereotype threat) of this thesis. Mindfulness involves effects on the body that derive from the embodied cognition approach. The following section will introduce embodiment, and later sections will cover the connection between embodiment and mindfulness and its effects on mental rotation performance.

2.7 Embodiment in Mental Rotation

2.7.1 Embodied Cognition

Embodiment theory emerged as a conceptual framework to understand the mind (Niedenthal et al., 2005). Before, traditional views on cognition were to assume that all individuals observe and represent the world in abstract symbols that they manipulate during a thinking process (Michalak et al., 2012). Opposed to this, Lakoff and Johnson (1999) described the brain taking input from the rest of the body. The idea is that cognitive processes are grounded in the organism's sensory and motor experiences through bodily interactions with the world, which lead to direct effects of bodily experiences on the mind (Barsalou, 2008). Therefore, embodiment is described as an essential part to explain the performance in cognitive tasks. The capability to skillfully adapt to varying circumstances (in sports, e.g., opponent's skill level, wet ground, wind) demonstrates the role of the body to include knowledge about the world (Khoury et al., 2017). An increasing amount of studies suggest that the body has an influence on perception and cognition (Chandler & Schwarz 2009; Jostmann et al., 2009; Natanzon & Ferguson, 2012; see Khoury et al., 2017). Wilson and Golonka (2013) further postulate that embodiment is an even more "radical hypothesis" (p. 1), and that the brain is not our only cognitive resource for problem solving. They elaborate that "our bodies and their perceptually guided motions through the world do much of the work required to achieve our goals, replacing the need for complex internal mental representations" (Wilson & Golonka, 2013, p. 1). Shapiro (2019) also reviewed the current state of embodied cognition and its future potential. However, he emphasized that embodied cognition could still be considered more as a "research program than a well-defined theory" (Shapiro, 2019, p. 2). He further states:

> In sum, the ties between subject matter, ontological commitments, and methodology are, within embodied cognition, longer and looser than they are within standard cognitive science. Yet, this state of affairs is no reason to dismiss or disparage embodied cognition. Today's research program may be tomorrow's reigning paradigm. However, embodied cognition's status as a research program does invite special caution when considering claims that it might replace or supersede standard cognitive science. (Shapiro, 2019, p. 3)

One way to go on with scientific endeavors is to focus on specific themes that are dominant in describing and illustrating embodied cognition, which offers a possibility to postpone the aforementioned "hard questions", until more is understood (Shapiro, 2019).

According to Shapiro (2019), future research will have to determine what embodied cognition is trying to explain, and then, whether embodied cognition and standard cognitive science have the same subject matter. Depending on this, science has to determine, if competing explanations of subject matters arise, and if so, whether only one or both of them are worth to pursue (Shapiro, 2019).

2.7.2 Embodied Stimuli in Chronometric Mental Rotation Tasks

An existing assumption is that pictures of abstract or non-human objects, like cube figures, are processed with an object-based mental transformation. In contrast, human body or body parts pictures are assumed to be embodied, and therefore, to be processed with an egocentric perspective-based mental transformation (Zacks & Tversky, 2005). However, as described in **Section 2.3.**, the stimulus type (embodied versus non-embodied) and task instruction type (egocentric versus object-based) may get confounded. But it has been shown that not the type of stimuli but the kind of instruction (left/right vs. mirrored/non-mirrored) determined, if an egocentric or object-based transformation is evoked (Voyer et al., 2017).

Based on the embodied cognition approach (see above, e.g., Lakoff & Johnson, 1999), mental rotation experiments use human body figures or body parts (e.g., pictures of hands) to evoke motor resonance processes, which would lead to better task performance by utilizing a sensorimotor simulation mechanism (e.g., Buccino et al., 2004; Calvo-Merino et al., 2005; Liuzza et al., 2012; Voyer & Jansen, 2016). This mechanism is also supported by behavioral and neuroimaging data indicating similar motor representations between observing, performing and mentally imagining an action (for reviews, see Decety, 2002; Dijkstra & Post, 2015). The findings of Amorim et al. (2006) indicate that the stimulus type is essential, due to increased familiarity, for instance, by adding a human head to cube figures. They further conclude that familiar postures would be easier to emulate than unfamiliar or atypical ones, eliciting embodied spatial transformations that facilitate task performance.

2.7.3 Behavioral Performance in Object-Based Transformations with Embodied Stimuli

In their neuroimaging study, Jordan et al. (2001) compared the brain mechanism and the performance in 3D cube figures with abstract shapes and letters. On the behavioral level, cube figures showed longer reaction times than the other two stimulus types. Similar to the study of Shepard and Metzler (1971), reaction times increased as a function of angular disparity. This is in line with the study of Campbell et al. (2018), who conducted an experiment with object-based transformations using Shepard and Metzler (1971) cube figures as well as human hand images. For all angles of rotation, the reaction times were higher in cube figures than for pictures of hands as stimuli.

As illustrated above, one important matter in mental rotation research is the potential existence of sex differences. Regarding the behavioral aspects, there is a continuous discussion about whether and how sex influences the performance in chronometric mental rotation tests, where the test is applied on a computer and reaction time and accuracy are measured (Jansen-Osmann & Heil, 2007). If differences between men and women exist, they seem to be partially explained by task complexity and stimulus dimensionality. In three-dimensional tasks, men outperform women, whereas in two-dimensional tasks no such difference is observed (Roberts & Bell, 2003). This might indicate that possible sex differences are generally reduced when objects are easier to process, which would also apply to embodied objects. Whereas Amorim et al. (2006) reported the processing of embodied objects in general being easier than for abstract ones, they did not investigate sex differences. This result has been confirmed by one study of Voyer and Jansen (2016), who also examined sex differences. However, the stimulus type still showed a pronounced effect for males, that is, males performing more accurately than females, suggesting that the use of embodied stimuli does not particularly favor females to perform better. Campbell et al. (2018) reported contradicting results using cube figures and pictures of hands as stimuli. Their findings did not show any differences on the behavioral level for cube figures, but did show females outperforming males in the mental rotation of pictures of human hands. Those studies demonstrate that the topic of sex differences in chronometric mental rotation tasks is still under discussion, and if they exist, the underlying mechanisms are not well understood. One possible explanation might be that females need more cognitive effort to solve object-based mental transformation tasks (Campbell et al., 2018). The mentioned effects and open questions were addressed in **Study 1** by using abstract and embodied figures as mental rotation stimuli, including possible sex differences, and analyzing cognitive effort. Two measurements of cognitive effort are highlighted in **Section 2.9** and form an important part of this thesis.

2.8 Mindfulness in Mental Rotation

2.8.1 Introduction to Mindfulness

The term *Mindfulness* has become popular over the last decades, which reflects the growing interest in mindfulness as a concept (Khoury et al., 2017). Mindfulness forms part of a variety of practices in the field of "meditation". Meditation techniques are often understood to try to elicit relaxation and, over a lengthy period of time, an increased sense of enduring happiness (Ekman et al., 2005) and well-being (Fischer, 1971; Lutz et al., 2008). However, this broad description can lead to problems in the attempt to scrutinize details of various meditation practices (Lutz et al., 2008). Lutz et al. (2007, 2008) proposed a more specific conceptualization describing meditation as "a family of complex emotional and attentional regulatory strategies developed for various ends, including the cultivation of well-being and emotional balance" (Lutz et al., 2008, p. 163). These intended effects could be observed in a lot of controlled studies, which indicated a relationship between mindfulness meditation and improved mental health regarding various disorders, such as anxiety disorders (e.g., Hofmann et al., 2010; King et al., 2013), depression (e.g., Coelho et al., 2007; Hofmann et al., 2010), eating disorders (e.g., Kristeller et al., 2006; Tapper et al., 2009), substance abuse (e.g., Bowen et al., 2006), and chronic pain-symptom reduction (Zeidan, Gordon, et al., 2010; for a review, see Wolkin, 2015).

The two main conceptualizations of *Mindfulness* are the traditional Buddhist approach and the contemporary western approach. As the term is used broadly, various approaches among different scholars apply different meanings to the same word (for a review and an extensive list of definitions, see Khoury et al., 2017). Buddhist mindfulness incorporates observing the emotional and cognitive processes, so that the mind shall not wander or become consumed by the future or the past (Lama & Berzin, 1997; Khoury et al., 2017). A newer definition also refers to embodiment by describing the mind, body, and external world interacting with each other (Stanley, 2013). As Khoury et al. (2017) point out, Buddhist mindfulness has not been investigated in empirical studies yet, although some studies regarding the western-based approach implemented parts of the ethical and spiritual elements of Buddhist mindfulness (e.g., in Avants et al., 2005; Rapgay & Bystrisky, 2009). Still, more research is needed to validate this traditional Buddhist mindfulness conceptualization to determine its effectiveness (Khoury et al., 2017). Regarding the western conceptualizations of mindfulness, Jon Kabat-Zinn and colleagues (e.g., Kabat-Zinn, 1982, 1991, 1994, 2003; Kabat-Zinn et al., 1985, 1986; Williams & Kabat-Zinn, 2011), as well as Ellen Langer and associates (e.g., Langer, 1989, 1997, 2009, 2011, 2012; Langer & Imber, 1979; Langer & Piper, 1987; Langer et al., 1978, 1989, 2009) developed two of the main approaches that are prominent nowadays. Based on the Buddhist approach, Kabat-Zinn described mindfulness-meditation as “paying attention in a particular way: on purpose, in the present moment, and non-judgmentally” (1994, p. 4). Later, this became one of the most used definitions of western mindfulness-meditation in scientific literature, and was adapted, modified, and expanded by various authors, focusing on and including aspects like attention, awareness, intention, attitude, acceptance, and non-judgment (for a detailed review, see Khoury et al., 2017). Khoury et al. (2017) point out that “directing attention” is overly prominent in all the definitions of mindfulness-meditation, and that it indicates the importance of self-regulation of attention to define and operationalize mindfulness. Over the following years, Kabat-Zinn developed different programs, which eventually led to the *Mindfulness-Based Stress Reduction* or MBSR. This protocol consists of practices from Buddhism (e.g., breathing, sitting, eating meditation, body scanning, and gentle stretching) and also western psychological methods (e.g., psycho-education, group discussions, and individual support; Khoury et al., 2017). Throughout the MBSR program, the body-oriented practices (e.g., body scanning and mindful eating) are often used, and further recommended as part of the home exercises. These MBSR programs gained more and more popularity, and an increasing amount of studies were conducted using Kabat-Zinn's or adapted protocols (e.g., *Mindfulness-Based Cognitive Therapy*, MBCT; Grossman et al., 2004; Segal et al., 2013; see Khoury et al., 2017). Analyzing 20 recent systematic reviews and meta-analyses, Gotink et al. (2015) reported positive effects of MBSR and MBCT programs on symptoms of depression, anxiety, and stress. Although the exact treatment mechanisms are still not fully known, a lot of authors remark an effective role of attention and emotional regulation (e.g., Chiesa et al., 2013; Hofmann et al., 2012; Hölzel et al., 2011; Nielsen & Kaszniak, 2006; Teasdale et al., 1995; Wolkin, 2015; also see Khoury et al., 2017).

Lutz et al. (2008) proposed a framework of operational definitions for meditative practices with the purpose that, within this framework, studies could examine the effects of meditation training on the mind and the brain. They focused on the meditation practices using Buddhist techniques and clinical secular derivatives (e.g., MBSR), and narrowed it down to two meditation styles: focused attention meditation and open monitoring meditation. Within one or over several sessions of training, these two meditation styles can also be combined (e.g., in the MBSR program).

Focused attention meditation implies directing and sustaining the focus of attention on a chosen object. For this, three main skills are being honed to regulate attention: staying vigilant to distractions while maintaining the intended focus; being able to disengage from a distraction; and being able to quickly redirect the focus on the selected object. Meditation practice shows progress partially through the amount of effort needed to direct and maintain the intended focus (Lutz et al., 2008). Advanced meditators notice faster, when the mind starts to wander, which eventually leads to a "trait change, whereby the attention rests more readily and stably on the chosen focus. At the most advanced levels, the regulative skills are invoked less and less frequently, and the ability to sustain focus thus becomes progressively *effortless*." (Lutz et al., 2008, p. 164).

Open monitoring meditation describes a non-reactive monitoring of all experiences from moment to moment. For this, it uses and builds on the focused attention training, with a specific focus on developing the monitoring skill (Lutz et al., 2008). With practice, "the *effortful* selection or *grasping* of an object as primary focus is gradually replaced by the *effortless* sustaining of an awareness without explicit selection" (Lutz et al., 2008, p. 164). In their review of the neuroscientific study of focused attention meditation, the authors also concluded that "at least several subcomponents of attention are best regarded as the product of trainable skills, and that focused attention meditation represents a family of mental practices that are explicitly designed to train such attentional skills" (Lutz et al., 2008, p. 165).

The concept of mindfulness by Langer et al. (1978) is different to the Buddhist conceptualizations and western derivatives. Langer's (1997) definition of mindfulness includes "being open to novelty, sensitive to context and perspective, challenging assumptions and predefined categories, and taking responsibility" (see Khoury et al., 2017, p. 1164). These definitions were later operationalized to define parts of Langer's concept like openness to novelty, flexible thinking, and cognitive reframing (Pagnini & Philips, 2015), with techniques such as "noticing distinctions, multiple perspectives, or producing novelty" (Alexander et al., 1989; see Khoury et al., 2017, p. 1164). Research utilizing Langerian mindfulness treatments showed effects on problem solving (Ostafin & Kassman, 2012), learning (Langer et al., 1989), performance (Langer et al., 2009), creativity (Grant et al., 2004), attention, and cognitive flexibility (Levy et al., 2001; for a review, see Khoury et al., 2017).

Longer mindfulness meditation practices over several weeks lead to an improvement in attentional control (Lutz et al., 2008), but also a short meditation practice can impact cognitive control tasks in a different manner depending on the kind of the short meditation form (Colzato et al., 2016). A brief mindfulness induction is often a short single session of mindfulness, for instance, to eat a raisin mindfully (Schofield et al., 2015; Weger et al., 2012).

To sum it up, the mechanisms of mindfulness can be described as the following: attention regulation (e.g., maintaining attention on a chosen object), emotion regulation (comprising aspects of reappraisal; exposure, extinction, or reconsolidation), body awareness (e.g., experiencing sensations of breathing and emotions), and a change in perspective (detachment from identification with a static sense of the self) (Hölzel et al., 2011; see Portele & Jansen, 2023).

2.8.2 Mindfulness and Embodiment

Within Buddhist mindfulness practices, the full awareness of all sensations occurring in every moment is a key element. The necessity to actively and intentionally direct attention to the processes of the body, feelings, mind, and phenomena defines this present moment awareness (Hart et al., 2013, Thera, 2005; see Khoury et al., 2017). Accordingly, mindfulness-meditation (Buddhist and western derivatives, such as focused attention meditation) can change formerly learned patterns of how the body, mind, and environment interact (Lutz et al., 2008; Wallace, 1999; see Khoury et al., 2017). Neuroscientific research indicated that, compared to non-meditation, MBSR training lead to higher activation levels in the brain regions, which are both related to interoceptive (e.g., insula) and exteroceptive (e.g., somatosensory cortex) body awareness (e.g., Chiesa et al., 2013; Hölzel et al., 2011; Lutz et al., 2008; see Khoury et al., 2017). In accordance to these findings are the descriptions of Niedenthal et al. (2005) of embodiment processes to be evident in both the peripheral (body-based) and central (modality-based) sense of the term embodiment. The brain's modality-specific systems include sensory systems (underlying perception), motor systems (underlying action), and introspective systems (underlying conscious experience of emotion, motivation, and cognitive operations). Furthermore, Varela and colleagues (e.g., Thompson & Varela, 2001; Varela et al., 1991) suggested consciousness to be embodied through a two-way reciprocal relationship between body and brain. This also connects to Cauller (1995), who described in his review, that top-down corticocortical (i.e., connecting one cortex to the other) influences interact with bottom-up sensory feedback in the primary sensory areas, which are central to conscious processing.

2.8.3 Measurements of Mindfulness

Over the last decades, several measurements of mindfulness were developed. They are normally self-reported multi-dimensional scales and may include various aspects of mindfulness meditation, such as attention, awareness, present focus, and acceptance/non-judgment. They can also vary regarding the measures of state and trait mindfulness (see Kiken et al., 2015; Lutz et al., 2008). Considering the embodied aspects in mindfulness meditation, Khoury et al. (2017) offered critique that western mindfulness-meditation scales often do not explicitly refer to the awareness of bodily sensations. For instance, they are not included in the often used Mindfulness Attention Awareness Scale (MAAS; Brown & Ryan 2003), which consists of 15 items with a 7-point scale, and is a self-report instrument measuring attention to and awareness across several areas of experience in daily life (e.g., cognitive, emotional, physical, and general). Scores of the MAAS strongly correlate with self-consciousness, rumination, and self-reflection. In other scales, body sensations are briefly mentioned, such as the Kentucky Inventory of Mindfulness Skills (KIMS; Baer et al., 2004), and even more so, in the State Mindfulness Scale (SMS; Tanay and Bernstein, 2013). The SMS does include the awareness of bodily sensations, but does not refer to the body-mind connection (see Khoury et al., 2017).

In order to measure the effects of short mindfulness interventions, not all scales are suitable, as some are more focused on measuring the development throughout the length of mindfulness programs. For single bouts of mindfulness interventions, the Toronto Mindfulness Scale (TMS; Lau et al., 2006) offers a possible measurement tool. It is a 13-item scale that is structured in two factors (Curiosity, Decentering). It is uniquely state-oriented for use immediately following a meditation experience, and has been validated in a number of clinical contexts (Lau et al., 2006; Davis et al., 2009). Lau et al. (2006) described the two factors as the following:

> The items of Factor 1 (Curiosity) reflect an attitude of wanting to learn more about one's experiences. The items of Factor 2 (Decentering) reflect a shift from identifying personally with thoughts and feelings to relating to one's experience in a wider field of awareness. (Lau et al., 2006, p. 1460-1461; see also Teasdale et al., 2002)

Regarding the aforementioned critique of Khoury et al. (2017), the TMS does also not include items referring specifically to the body mind connection. However, it offers the mentioned possibility to measure state mindfulness through focused attention meditation.

2.8.4 Mindfulness in Combination with Stereotype Threat Induction

An important research question is how potential stereotype threat effects can be reduced. In one study with a mathematical task, Weger et al. (2012) used a brief mindfulness intervention to possibly alleviate the effects of stereotype threat. 71 female psychology students (mean age = 20.14 years, SD = 3.99) participated, who had to complete two mathematical tests with or without a mindfulness intervention, and with or without stereotype activation. The authors reported a significant main effect of a five-minute mindfulness intervention (eating two raisins) versus a control condition ($F(1,67) = 5.61$, $p = .021$, $\eta^2 = .077$). The results also showed that performance decrements that typically occur under stereotype threat could be reversed after the short mindfulness induction.

In line with Weger et al. (2012), Rahe and Jansen (2023) reported a significant effect of a five-minute mindfulness intervention (eating two candies) in their group of 152 adolescents (80 boys, mean age = 16.59 years, SD = 0.780, and 72 girls, mean age = 16.25 years, SD = 0.904;

$t(150) = 4.768$, $p < 0.001$, $d = 0.779$) in their recent study with psychometric mental rotation tasks. However, they found no significant stereotype activation.

To our best knowledge, only one study for chronometric mental rotation exists to date, where it was shown that participants of both sexes responded faster after mindful learning (Geng et al., 2011). With an overall sample of 32 students, they found significant performance improvements due to mindful learning with no significant differences between men and women. However, the study did not include a stereotype induction.

Apart from the main effect of mindfulness mentioned above, Weger et al. (2012) reported an interesting interaction between stereotype threat and mindfulness intervention, that is, that the math performance was only ameliorated in a stereotype threat condition with a mindfulness induction before. The authors explained it with a disassociation of the cues linked to social comparison from their threatening value, which would lead to a reinterpretation of the threat because of the brief mindfulness intervention. At least three explanations for this result could be considered: One of them is that mindfulness could improve emotion regulation (Heppner et al., 2008), which is debilitated by stereotype threat (Johns et al., 2008; see also Weger et al., 2012). Another explanation is that a short mindful induction influences working memory, which might be worsened in stereotype threat situations (e.g., Beilock et al., 2007). Furthermore, one can assume that stereotype threat takes up cognitive resources and increases cognitive effort. Lastly, mindfulness, emotion regulation, working memory, and stereotype threat could be linked together, as emotion regulation can influence working memory and vice versa (see Barkus, 2020; Groves et al., 2020). For these reasons, we wanted to investigate the effects of mindfulness on cognitive effort with and without stereotype threat induction. The next section addresses the possibilities to measure cognitive effort.

2.9 Measurements of Cognitive Effort

2.9.1 Pupillometry as a Measure of Cognitive Effort

Pupil dilation has been investigated since the 1960s (e.g., Hess & Polt, 1964; Kahneman & Beatty, 1966), but has become more popular again during the last two decades, thanks to advancements in computer and eye tracker technology (van der Wel & van Steenbergen, 2018).

A large variety of cognitive processes can influence the eye's pupil size (for reviews, see Beatty & Lucero-Wagoner, 2000; Loewenfeld, 1958; Mathôt, 2018; van der Wel & van Steenbergen, 2018). Three mature types of stimuli can cause pupil size changes: constriction due to brightness (pupil light response), near fixation (pupil near response), and dilation due to an increase in cognitive activity like increased levels of arousal or mental effort (psychosensory pupil response; Mathôt, 2018).

Regarding the pupillary light response, an increase of brightness leads to pupil constriction and even occurs, when the source of brightness is only imagined (Laeng & Sulutvedt, 2014), covertly attended to (Binda et al., 2013; Mathôt et al., 2013; Naber et al., 2013; Unsworth & Robison, 2016), read about (Mathôt et al., 2017), or maintained in visual working memory (Hustá et al., 2019; Zokaei et al., 2019; see Mathôt & Vilotijević, 2022). The pupil near response involves a reflective constriction of the pupil in response to looking at a nearby object, and, vice versa, pupil dilation by looking at a far-away object. The psychosensory pupil responses include an orienting response (brief involuntary pupil dilation as a consequence of directing attention to something) and slower arousal- or mental-effort-related responses related to high-level cognition (Mathôt, 2018). Regarding the latter, Kahneman (1973) described an exercise to easily experience the connection between mental effort and pupil dilation:

> Face a mirror, look at your eyes and invent a mathematical problem, such as 81 times 17. Try to solve the problem and watch your pupil at the same time, a rather difficult exercise in divided attention. After a few attempts, almost everyone is

> able to observe the pupillary dilation that accompanies mental effort, in a situation which elicits neither overt responses nor test anxiety. (Kahneman, 1973, p. 24)

In his review, Mathôt (2018) describes the neurological pathways related to pupil dilation. The iris sphincter muscle is innervated by the parasympathetic nervous system (the part of the autonomic nervous system that keeps the body in a stable condition), which is the reason why pupils are relatively small at rest. This muscle is also connected to the retina via the constriction pathway, which is a subcortical pathway and considered crucial for pupil constriction, that is, mainly through the pupil light response (Mathôt, 2018). Accordingly, its counterpart, the iris dilator muscle, is connected to the dilation pathway, which is controlled by the sympathetic nervous system (the part of the autonomic nervous system that is linked to arousal, wakefulness, and the fight-or-flight response). The dilation pathway is a subcortical pathway, which begins at the hypothalamus and the locus coeruleus, and connects to the iris dilator muscle (Mathôt, 2018). The locus coeruleus is active in an organism's aroused, awake, and alert state. The constriction and dilation pathways are generally distinct, but also interact in several ways. Most notably, the locus coeruleus inhibits the parasympathetic constriction pathway at the level of the Edinger-Westphal nucleus (which receives input from the left and right pretectal nucleus, and combines the information of both visual fields). More specifically, iris dilation is regulated by the locus coeruleus-norepinephrine system mainly via norepinephrine stimulating α-adrenoceptors of the iris dilator muscle, and postsynaptic α_2-adrenoceptors of the Edinger-Westphal nucleus that projects to the ciliary ganglion (Yoshitomi et al., 1985). Since these dilation adaptations are completely different to contractions due to the pupillary light reflex (via acetylcholine), constant low light levels are critical to reliably measure norepinephrine levels (Aston-Jones & Cohen, 2005; Koss, 1986; Nieuwenhuis et al., 2005). This described pathway was suggested to primarily underlie pupil dilation caused by arousal and mental effort (Steinhauer et al., 2004; see Mathôt, 2018).

The field of *Cognitive Pupillometry* describes the measurement of pupil size to investigate cognitive processes (Mathôt & Vilotijević, 2022). As introduced above, in the research studies presented here, the focus is on the information pupillometry can give us regarding cognitive effort. The literature regarding the analysis of changes of pupil diameters is vast. Applying eye-tracking technology is a difficult endeavor, and for the pupillometric data to be valid, a lot has to be considered, such as laboratory setup (e.g., room lighting, monitor viewing distance and angle), data capturing hardware and software (e.g., eye-tracker type and specifications, capture software with different algorithms), pre-processing (e.g., parsing raw data, interpolating or removing missing and invalid data, baseline correction, trial and participant exclusion based on data quality), and finally statistical analysis (e.g., single test, cluster-based permutation, cross-validation; see e.g., Holmqvist et al., 2011; Mathôt, 2013, 2018; Mathôt et al., 2018; Mathôt & Vilotijević, 2022).

Thus, if everything is controlled for, cognitive pupillometry describes the measurement of the rapidly changing pupil diameter during cognitive processes. When set in relation to baseline values, the task-evoked changes of the pupil diameter do normally not exceed 0.5 mm (Beatty & Lucero-Wagoner, 2000), and can be used as a "psychophysiological index or correlate of cognitive activity" (Campbell et al., 2018, p. 20) that changes because of task difficulty (Hess & Polt, 1964; Kahneman & Beatty, 1966). In all cognitive tasks where behavior is compared for conditions of varying task difficulty, van der Wel and van Steenbergen (2018) describe a persisting problem when looking at pupil dilation as an indicator for neural and computational mechanisms underlying cognitive effort. According to the authors, when analyzing the mean differences of such conditions, no easy decision can be made whether the physiological signals indicate mere task demand or effort exertion. In their review, they describe and discuss evidence and arguments in favor for both interpretations. Based on the overall findings, they conclude that "the effort account could have more explanatory power and might therefore be preferred over an account that pupil dilation reflects mere task demands“ (van der Wel & van Steenbergen, 2018, p. 2010).

Kahneman (1973) used the terms *capacity*, *effort*, and *attention* interchangeably to describe the limited working memory resources available to participants while solving cognitive tasks. Similarly, a lot of studies (e.g., reviewed in Beatty & Lucero-Wagoner, 2000; Mathôt, 2018) showed pupil size to reflect mental effort, cognitive load, or cognitive intensity. In summary, the pupils dilate as a consequence of whatever activates the mind (Mathôt, 2018).

One influential theory that was first proposed by Aston-Jones and Cohen (2005) is the adaptive-gain theory. It primarily involves how the locus coeruleus regulates behavior. As mentioned above, the activity in the locus coeruleus has a high correlation with pupil size (Joshi et al., 2016). For that reason, its measurement via pupillometry can be used to test the predictions of the adaptive-gain theory (e.g., Gilzenrat et al., 2010; Jepma & Nieuwenhuis, 2011; for a review, see Mathôt, 2018). Aston-Jones and Cohen (2005) describe two distinct modes of behavior, which they refer to as *exploitation* and *exploration*. Exploitation describes a behavioral mode where one is engaging in one single task and exploiting the rewards of this task, for instance, eating food. The mode is linked to intermediate, phasic activity in the locus coeruleus and thus intermediate pupil size (see Mathôt, 2018). The other mode, exploration, describes a state where one can switch between different tasks and can also be distracted more easily. In this state, one is exploring, which one of possible tasks offers the highest reward. Animals and humans show behaviors that coincide with this theory, as switching between exploration and exploitation can be observed. Mathôt (2018) pointed out that the tracking of this switching behavior with pupillometry is a new interesting approach and that there are some studies with supporting results. One example is a study by Jepma and Nieuwenhuis (2011), who reported pupils being larger in exploration than exploitation. Additionally, stimuli elicited large but less responsive pupil responses in exploration (Gilzenrat et al., 2010). These two aspects of the exploration pattern reflect neural activity of the locus coeruleus (Aston-Jones & Cohen, 2005; see Mathôt, 2018).

2.9.2 Subjective Cognitive Effort

Cognitive pupillometry is one way to objectively measure changes in mental effort and cognitive load. However, this method cannot be used in all experiments, depending on variables such as test design, apparatus availability, and other constrictions. To ascertain cognitive load demands of test problems in these type of experiments, self-rating scales of mental effort can be used (see Paas & van Merriënboer, 1994; Paas et al., 2003). Normally, they are Likert-type scales, consisting of a nine- or ten-point scale. For instance, Ayres (2006) used a nine-point scale: 1 (extremely easy), 2 (very easy), 3 (easy), 4 (quite easy), 5 (neither easy or difficult), 6 (quite difficult), 7 (difficult), 8 (very difficult), and 9 (extremely difficult). A similar and German version is the "Anstrengungsskala Sport" (ASS), which is an RPE (rate of perceived exertion) scale developed by Büsch et al. (2015) ranging from 0-10. In their study regarding mental rotation and aerobic exercise, Jost et al. (2023) used this scale to measure the RPE for both subjective physical exertion and subjective cognitive exertion. Measurement with the scale is straightforward: Participants have to name or mark a unique number or word to describe the perceived cognitive effort. Depending on the test design, these measurements can be done after each task or after whole test blocks. Mental rotation experiments usually include an extensive number of items to ensure test power. Including a subjective rating of perceived exertion after each item would drastically increase the experiment duration, which has to be taken into consideration. For instance, Jost et al. (2023) measured the subjective cognitive effort after each mental rotation task block and analyzed these block values. Similarly, Ayres (2006) measured cognitive load once after finishing each whole answer booklet in his study on mathematical learning.

2.9.3 Cognitive Effort in Mental Rotation Tasks

Until now, to our knowledge, only one study investigated object-based mental rotation tasks using different stimulus types (embodied and abstract) measuring cognitive effort with respect to sex differences: Campbell et al. (2018) implemented the physiological correlate of cognitive effort in a mental rotation task design. They compared abstract (cube) with embodied (human hand) figures in chronometric mental rotation tasks with 50 males and 49 females applying pupillometry. Their findings confirmed that pupil dilation was modulated by angular disparity, with higher angular disparity, that is, more difficult tasks, increasing the pupil diameter. Additionally, females showed higher cognitive load than males in mental rotation tasks of abstract figures. For the embodied figures, for which they used hand pictures, both sexes showed comparable levels of cognitive effort, indicating that due to embodiment, sex differences in this spatial task dissipated. However, in the study of Campbell et al. (2018), the embodied stimuli (human hands) and the cube figures differed completely in number and kind of visual features, because they did not have any shape or color features in common. A comparison might therefore be more difficult. Another study that was mentioned above is from Jost et al. (2023). As one of the findings in their study regarding mental rotation and aerobic exercise, they reported subjective cognitive effort being significantly lower for women than for men.

Apart from task difficulty, monitoring the pupil diameter might illustrate other effects. One might be the effect of stereotype threat to deplete working memory capacity (Spencer et al., 2016). Working memory capacity is connected to cognitive tasks such as reasoning (Suess et al., 2002) and inhibitory control (Redick et al., 2011). Considering the various processing stages in mental rotation (encoding, identification, rotation of the stimuli, the response selection and motor response; Heil & Rolke, 2002), some parts of these, such as stimulus encoding and object comparison, can reflect functions of the working memory, such as processing of visual or object working memory (Hyun & Luck, 2007). Research using fMRI has shown that in adults, the mental rotation process involves the frontal cortex (BA , BA10), premotor cortex (BA6), and parietal cortex (BA 40, BA 44) (Cohen et al., 1996; Jordan et al., 2001; Schöning et al., 2007; see Yang et al., 2020). Thus, the working memory seems to play an important role while solving mental rotation tasks (e.g., Hyun & Luck, 2007; Kaufman, 2007; Pardo-Vazquez & Fernandez-Rey, 2012). Booth et al. (2000) also specified that mentally rotated stimuli are temporarily stored in the working memory. This was supported by Gathercole et al. (2004), who described the visuospatial sketchpad—a subsystem of the working memory—playing an important role in the manipulation of visual images. The role of this subsystem is also supported by the findings of Lehmann et al. (2014), who reported a positive correlation between spatial working memory (operationalized by a block-tapping-test) and mental rotation capacity. In addition to the behavioral findings, similarities in neurological activation support this notion of a connection between working memory and mental rotation. Anguera et al. (2010) found activity in the dorsolateral prefrontal cortex and the bilateral inferior parietal lobule while solving visuospatial working memory tasks. These areas were also reported to be active during the early processes of mental rotation (Cohen et al., 1996; Jordan et al., 2001). The working memory is usually described as a capacity-limited system to store and manipulate information during a short time period (e.g., Robison & Unsworth, 2019). However, instead of a single storage, it was suggested to be better considered as a "model that consists of separate domain-specific components that interact with each other" (Lehmann et al., 2014, p. 561). One model of Baddeley (1992, 2000) described the working memory to consist of four subcomponents: the central executive, the episodic buffer, and the two subsystems of the phonological loop and the aforementioned visuospatial sketchpad. Although parts of these components have been shown to be involved in mental rotation processes (e.g., Hyun & Luck, 2007; Lehmann et al., 2014), Lehmann et al. (2014) remarked that it is still unclear whether all parts of working memory are involved in mental rotation.

Generally, some studies indicated that higher working memory capacity is linked to better mental rotation performance (Hyun & Luck, 2007; Kaufman, 2007; Pardo-Vazquez & Fernandez-Rey, 2012). Interestingly, Goodmon et al. (2019) reported higher working memory capacity participants not being faster than lower working memory capacity ones in mental rotation. They also pointed out that only the latter showed the typical reaction time increase depending on angular disparity. They concluded that people with higher working memory capacity were not impeded in their performance as much as those with lower working memory capacity, and that this might also indicate different strategy approaches between these groups.

However, they pointed out an important limitation in their study that they only used reaction times as the dependent variable and excluded accuracy.

Robison and Unsworth (2019) examined fluctuations in working memory performance. They found that the variability in pretrial pupil diameter was the strongest independent predictor of individual differences in task performance. Their former studies had shown similar results, that is, individuals, who have more fluctuations in arousal, showed worse task performance, and those, who used more attention on maintaining the items during delay intervals, showed better task performance (Unsworth & Robison, 2015, 2018; see Robison & Unsworth, 2019). These findings support their theory of individual differences in how the locus coeruleus-norepinephrine system functions with regard to working memory and attention (Unsworth & Robison, 2017a, b). This theory included:

> Stability and consistency of delivery of NE [norepinephrine] to the frontal cortices by the LC [locus coeruleus] is an important underlying factor in why people vary in their ability to control their attention and effectively utilize their limited WM [working memory] system. When NE [norepinephrine] is not consistently delivered, arousal tends to fluctuate, which can lead to lapses of attention and be harmful for the performance of tasks that require consistent attention. (Robison & Unsworth, 2019, p. 408; see also Unsworth & Robison, 2017a, b)

Overall, pupillometry is an important measure of biomarkers for arousal, attention, and general task management (Clewett et al., 2020; Eldar et al., 2013; Gilzenrat et al., 2010; Joshi et al., 2016; Kahneman & Beatty, 1966; see Keene et al., 2022). More specifically, pupil diameter changes were shown to be connected to sustained attention (Decker et al., 2020; Unsworth & Robison, 2016; van den Brink et al., 2016) and working memory tasks (Robison & Brewer, 2020; Robison & Unsworth, 2019; Unsworth & Robison, 2015; Zokaei et al., 2019; see Keene et al., 2022).

3 Summary of the State of Research

Over the last five decades, mental rotation has been an important part of research in the field of spatial abilities. Because of the practical implications of the connection between spatial abilities and success in STEM disciplines, it continues to be an important topic and a popular field (e.g., Xie et al., 2020). In object-based mental rotation transformation tasks, participants have to decide whether two objects, which are rotated to each other, are mirror reversed to each other or not. In the analysis of chronometric mental rotation tasks, the behavioral data (reaction times and accuracy rates) show a generally linear relationship with the increasing angular disparity between the two figures (Shepard & Metzler; 1971).

Mental rotation ability can be influenced by many factors, of which stereotype threat and embodiment effects are highlighted in this thesis. Stereotype threat describes the risk of confirming a negative stereotype about the *own* group (Steele & Aronson, 1995), which may trigger a disruptive state that can reduce the performance, for example, in cognitive tests (e.g., Spencer et al., 2016). The effectiveness of stereotype threat depends on various factors; for instance, stronger beliefs about one's ability may reduce such effects (Rahe & Jansen, 2023). One possibility to change the beliefs in the own ability and a way to mitigate the effects of stereotype threat is the practice of mindfulness.

Mindfulness involves effects on the body that derive from the embodied cognition approach. The idea is that cognitive processes are grounded in the organism's sensory and motor experiences through bodily interactions with the world, which lead to direct effects of bodily experiences on the mind (Barsalou, 2008). Hence, the brain takes input from the rest of the body (Lakoff & Johnson, 1999), which can have an influence on perception and cognition (for a review, see Khoury et al., 2017). Longer mindfulness meditation practices over several weeks lead to an improvement in attentional control (Lutz et al., 2008) and emotion regulation (Heppner et al., 2008), but also short embodied meditation practice (e.g., eating a raisin mindfully) may impact cognitive control tasks (Colzato et al., 2016), for instance, by influencing the working memory, which might be worsened in stereotype threat situations (e.g., Beilock et al., 2007).

While embodiment effects can be elicited via external interventions such as meditation, cognitive tests by themselves may incorporate aspects that can provoke embodiment effect. For example, the figures serving as stimulus material may be abstract such as cube figures, or embodied such as figures of human bodies (Amorim et al., 2006) or hands (Campbell et al., 2018).

Another important role play the large sex differences discovered in psychometric mental rotation tests, which are not as easily replicated and observed in chronometric mental rotation tests. As there is yet much to discover, cognitive psychology still scrutinizes gender differences with new analytical methods, advancing technology, and with regard to various possible factors and indicators of cognitive processes. One is Cognitive Pupillometry, the measurement of pupil size to investigate cognitive processes (Mathôt & Vilotijević, 2022). By controlling various factors, changes in pupil dilation can serve as an objective biological marker of cognitive effort that is connected to the locus coeruleus-norepinephrine system (see Mathôt, 2018).

The following studies endeavor to address parts of the mentioned aspects. With regard to possible gender differences, three mental rotation experiments were conducted that included influences of embodied cognition and indicators of cognitive effort, and were analyzed with linear mixed models. As an objective biological marker of cognitive effort, changes in pupil size are recorded via eye-tracking technology. Embodied stimuli as well as treatment through mindfulness delve into the effects of embodiment and focused attention meditation. Lastly, the effects of the explicit emotional influence of gender stereotypes are investigated.

In **Study 1**, we analyzed mental rotation task performance with three types of stimuli that had different levels of embodiment, and gender. Here, pupillometric measurements were taken with eye-tracking technology to analyze cognitive load. In **Study 2**, the effects of a 20-minute embodied focused attention mindfulness meditation were analyzed. The meditation and a control treatment, and gender were investigated regarding mental rotation performance and subjective cognitive effort. In **Study 3**, we used a shorter embodied focused attention mindfulness meditation than in the second study, and combined it with explicit stereotype threat applied through auditory induction. All conditions and gender were investigated regarding mental rotation performance and pupillometric measurements of cognitive effort.

Throughout the studies, we endeavored to implement statistical advancements such as linear mixed models to improve the power of and the confidence in the results. For the pupillometric measurements, we prepared the laboratory conditions and conducted all measurements and data processing according to current standards, professional advice and guidelines, and to the best of our knowledge.

4 First Study: Pupillometry as a Measure of Cognitive Load in Mental Rotation Tasks with Abstract and Embodied Figures[1]

4.1 Summary

In the first study, we investigated sex differences in behavioral performance and cognitive load in chronometric mental rotation tasks with abstract and embodied figures. Eighty participants (44 women and 36 men) completed 126 items each, which included cube figures, body postures, and human figures, which were all comparable in shape and color. Reaction time, accuracy, and cognitive load, measured by changes in pupil dilation, were analyzed. As a function of angular disparity, participants showed shorter reaction times and higher accuracy rates for embodied stimuli than cube figures. Changes in pupil dilation showed a similar pattern, indicating that mental rotation of embodied figures caused less cognitive load to solve the task. No significant sex differences appeared in any of the measurements.

4.2 Goals and Hypotheses

To exclude the effect of different stimulus features, we used the stimuli from Amorim et al. (2006). They created two kinds of human body figures as altered versions of the abstract cube figures. These figures differ in their similarity to the original cube figures and are labeled as body postures and human figures (Jansen et al., 2012). Using pupillometry, we also examined whether any differences between the stimulus types are influenced by sex, regarding cognitive load. The following hypotheses were investigated.

Sex differences in reaction time and accuracy have to be investigated. In line with Jansen-Osmann and Heil (2007), no sex differences in the behavioral data could be expected, however,

---

[1] The results presented in this chapter were published in advance in: Bauer, R., Jost, L., Günther, B., & Jansen, P. (2022). Pupillometry as a measure of cognitive load in mental rotation tasks with abstract and embodied figures. *Psychological Research, 86*(5), 1382-1396.

with respect to the study of Voyer and Jansen (2016), men might outperform women (Hypothesis 1).

We predicted behavioral task performance to be better with embodied figures, due to the familiarity and sensorimotor functions associated with them compared to cube figures for both sexes (Hypothesis 2; Amorim et al., 2006).

According to Campbell et al. (2018), we predicted stimulus type and angular disparity to influence the pupil dilation, with pupil dilation increasing as angular disparity increases, and generally being larger for cube figures. No differences between the two embodied figures were expected. We also expected reaction time to influence the changes of the pupil diameter due to the connection between task difficulty and longer response times (Hypothesis 3).

In line with the findings of Campbell et al. (2018), we expected sex differences in pupillometric data for the abstract (cube) figures, but not for the embodied figures, with women showing higher levels in cognitive load (Hypothesis 4).

4.3 Methods

4.3.1 Participants

In total, 109 students (60 women) participated in the study and received study credits. No monetary compensation was involved in this study. For 29 participants, the Software Development Kit (SDK) had disrupted the connection to the eye tracker and thus terminated the experiment, which led to an exclusion from analysis due to software failure. As a result, 80 students (44 women, mean age (SD) = 20.6 (2.1) years; 36 men, 21.9 (2.8) years) form the sample for statistical analysis. To take part in the study, participants needed unrestricted eyesight at close range or corrected eyesight with contact lenses. All participants were free from eye injuries and reported no relevant physical or mental limitations. Informed consent was obtained from all individual participants included in the study. The experiment was conducted according to the ethical declaration of Helsinki. Ethical approval for this study was not required in accordance with the conditions outlined by the German Research Society (DFG), where research that carries no additional risk beyond daily activities does not require Research Ethics Board Approval. We communicated all considerations necessary to assess the question of ethical legitimacy of the study.

4.3.2 Setup

Stimulus presentation and response handling were controlled with Presentation® software (Version 20.1 Build 12.04.17, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com) on a Dell Latitude E5540 Laptop, 14", 1366 x 768 px, 60 Hz. Below the bottom screen border, a RED250mobile (SensoMotoric Instruments GmbH, 2017) eye tracker, 250 Hz, was applied. Using iViewRED software (Version 4.4.26.0, SensoMotoric Instruments GmbH, 2017), all screen properties and the position of the tracker relative to the screen were integrated. With the iViewX_SDK (Version 4.4.10.0, SensoMotoric Instruments GmbH, 2017), the Presentation® script conducted a 13-point calibration at the start of each run. Calibration accuracy in form of vertical and horizontal dispersion was reported to be lower than 0.3° for all participants. The iViewRED software depicts the distance of the eyes to the screen. All participants were seated close to the table. Then, the laptop was positioned with the screen being at 60 cm from participants' eyes. No chinrest was used in this experiment. All participants were instructed to remain relaxed and to move as little as possible throughout the experiment. As a consequence, their position in the headbox of the eye tracker was always given.

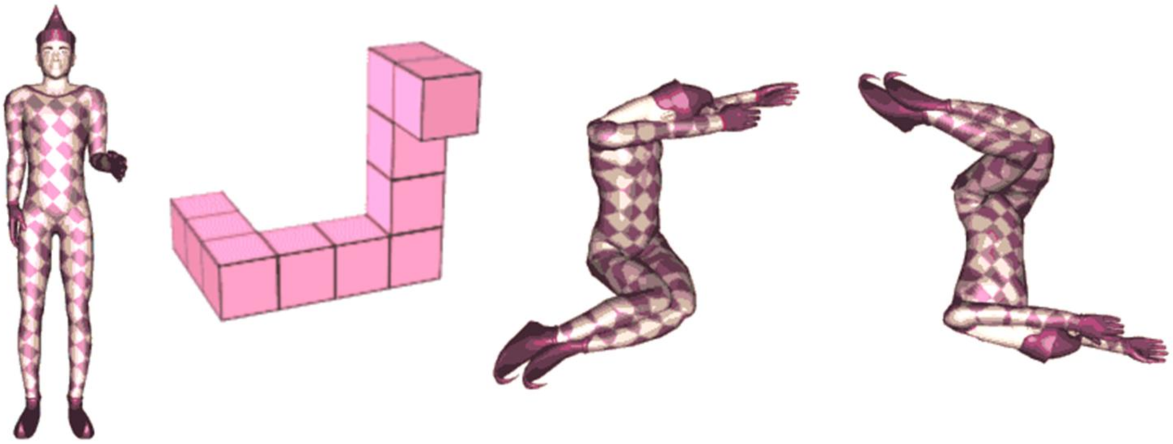
The eye tracker was placed on a table, which neither the participant nor the investigator touched during the experiment to prevent vibrating the tracker. To maintain an appropriate distance to the participant and to manage the software, the investigator used a wireless keyboard and mouse on a separate table. The participant used a wired mouse placed on a lower table beneath the tracker table to control for input lag and vibrations. The laboratory was silent and constantly dimly lit to control for pupil dilation due to light conditions. The only light sources were a ceiling lamp outside the peripheral visual field of the participant and the laptop screen (luminance at 169 cd/m^2 for body postures, and 177 cd/m^2 for both cube and human figures; measured with a spot meter, Chroma Meter CS-100, Minolta Co., Ltd., Japan), resulting in a constant illuminance of 55 lx (measured with a lux meter, testo 540, Testo AG, Germany).

4.3.3 Stimuli

Stimuli consisted of three different types developed by Amorim et al. (2006), which are Shepard and Metzler (1971) style 3D cube figures, human figures, and body postures (see Figure 2). Human figures are human bodies in standing positions holding their arms in different positions, matching familiar postures (e.g., shaking hands). Body postures are human bodies, whose atypical postures aligned with the cube figure configurations. All stimulus types had three variations each, were partially colored in pink, and were displayed in front of a white background.

Figure 2

Examples of Mental Rotation Stimuli Used in the Experiment.



Note. Developed by Amorim et al. (2006); reprinted and adapted with permission by Michel-Ange Amorim. From left to right: Human figure (non-mirrored). Cube figure (non-mirrored). Body posture (non-mirrored). Body posture (mirrored and rotated by 180°).

The participants looked at two three-dimensional figures (pairwise) and had to decide whether the figures were the same or different (mirrored). Each figure type was presented in a separate block (three blocks in total) with 42 trials each (total of 126 randomized trials) with one half of the trials being identical pairs and the other half mirror-reversed pairs. On the left side of the screen, the model was always presented non-mirrored with 0° rotation. On the right side of the screen a rotated and mirrored/non-mirrored stimulus was presented. The stimuli were presented in seven different angular disparities of 0°, 30°, 60°, 90°, 120°, 150° and 180° in y-axis (screen plane). Each figure had a dimension of 400 × 400 px and was vertically centered and horizontally positioned 300 px to the left or right of the center of the screen until a response was given. A practice block of 36 trials with feedback preceded the main experiment. Between stimuli pairs in the practice session, participants received feedback for 1000 ms (+ right, - wrong) shown at the center of the screen, and in experimental sessions, a fixation cross ("*") was shown there for 1000 ms. During the main experiment, self-controlled pauses were provided after every 14 trials.

4.3.4 Pupil Diameter

Before the practice block, the participants saw each stimulus pair of all nine figures in randomized order and only for the 0° non-mirrored condition. Participants were instructed that they were about to see some pictures, which they only had to look at. Originally, this first block was supposed to serve as baseline measurements, as Campbell et al. (2018) used them, for instance. However, we followed the recommendations by Mathôt and colleagues (Mathôt, 2013; Mathôt et al., 2018; Mathôt et al., 2015), and used a trial-dependent baseline correction. The response in pupil dilation typically occurs within the first few hundred milliseconds after stimulus onset, peaks around one to two seconds after stimulus onset and continues asymptotically until it returns back to baseline values (Andreassi, 2000; Beatty & Lucero-Wagoner, 2000; Loewenstein & Loewenfeld, 1962; Nieuwenhuis et al., 2011). More specifically, task-evoked pupil responses do not emerge earlier than ca. 220 ms after a manipulation that caused them (e.g., Mathôt et al., 2015). This opens up the possibility to take a period at the start of each trial for baseline correction. For this mental rotation experiment, we followed the approach of Mathôt et al. (2018) and used the median pupil size during the first eleven samples (corresponding to 40 ms), which was then used for subtractive baseline correction for each trial. Subtractive baseline correction is favorable, because it is more robust and increases statistical power more than divisive correction (see Mathôt et al., 2018, for an in-depth comparison). Overall, the baseline correction served to decrease the impact of random pupil-size fluctuations from one trial to the next, whereas between subject differences were taken into account statistically by using by-participant random intercepts in the linear mixed models (see Baayen et al., 2008; Mathôt et al., 2018).

Statistical inspection of the baseline data showed no significant differences between the angles (p = .073), but significant differences between the stimulus types (p = .007). Pairwise comparisons indicated that the pupil diameters for body postures were larger than for cube and human figures (which did not differ significantly from each other), illustrating the luminance differences between those stimuli.

4.3.5 Procedure

The experiment was a single session and lasted between 35 – 50 minutes, depending on participants' speed to complete all items. Upon arrival, the participants read and signed the informed consent. After that, they filled out a questionnaire including demographic information, sports activity, physical and mental illnesses, and eye-sight specifications. Then, they were positioned, as was the laptop respectively. After a brief explanation of the test protocol, the calibration and first block (presenting nine non-mirrored and non-rotated stimulus pairs for 6 seconds each) were run. Then the practice session with feedback followed, which was introduced by a digitally presented instruction. Participants used the right hand for mouse handling and received written instructions to press the left mouse button, if the stimuli could be rotated into congruence (non-mirrored, "same"), and the right mouse button, if the two stimuli were mirrored ("different"), and to answer as quickly and precisely as possible. Here, verbal feedback was only given, when participants did not understand the task, hence overly making mistakes or taking very long to respond (> 15 sec). After completing all practice trials, the calibration and main session were run. During the practice and main sessions, participants were asked to remain with their gaze on the screen (being allowed to blink naturally). During the short self-controlled breaks, they could avert or close their eyes for a few seconds. These instructions were necessary, because the SDK would produce an error and shut down the experiment upon longer gaze losses. Following the main experiment, the participants were debriefed. All verbal instructions were standardized using a researchers' guideline script. Three investigators conducted the data collection.

4.3.6 Study Design

To analyze cognitive performance, the dependent variables are reaction time (RT), accuracy (ACC) as well as the difference between the maximum pupil diameter and the respective baseline pupil diameter for each trial (PD) as a measure of cognitive load. To test our hypotheses, the independent variables are stimulus type (STI; cube figures [CF], human figures [HF], and body postures [BP]), SEX, angular disparity (DEG), and their respective interactions. DEG describes the angular disparity between the two figures shown on the screen. Since the left image is always presented with 0° rotation, the angular disparity depicts the rotation in degrees of the right image. DEG was included as a fixed effect as it is the main moderator of difficulty in mental rotation tasks (Jost & Jansen, 2020). In the analysis of PD, we included RT as a fixed effect to analyze the influence of reaction time on cognitive load.

4.3.7 Data Processing

For the behavioral data, outliers were determined by a deviance of more than three standard deviations from the mean reaction time of all stimulus pairs with the same rotation angle and were excluded from all analyses. Angular disparity is not unambiguously defined for mirrored responses in cube figures as they cannot be brought into congruence with each other (Jolicœur et al., 1985; Shepard & Metzler, 1971). Although mirrored figures have a unique shape, a mirror image may appear at different rotated positions due to different orientations of the mirroring plane. Thus, the angle between two mirrored images might be twice the one between the respectively rotated mirroring planes. As such, mirrored stimuli cannot have a unique 0°-condition and therefore no definable angular disparity (Jansen et al., 2020; see also Jost & Jansen, 2020). Because of that, only non-mirrored stimulus pairs were analyzed and reaction time was additionally only analyzed for correct responses. Using the SMI software BeGaze 3.7, build 58 (SensoMotoric Instruments GmbH, 2017), a velocity dependent algorithm (peak velocity threshold = 40°/s, min. fixation duration = 50 ms, peak velocity between 20–80% of saccade length) was used for blink detection. Here, blinks are saccade-like events during which fast pupil diameter changes occur (which reflects a rapid shrinking of the pupil due to closing of the eyelid). In R (R Core Team, 2018), the exported raw data (with marked blinks) were further processed to obtain valid pupil diameter data. The data for each eye were treated separately. As recommended by Mathôt et al. (2018), in addition to filtering based on detected blinks, we also filtered based on pupil size. A band pass filter was used to reject pupil size samples outside a predefined range between 1.5 and 9 mm (Kret & Sjak-Shie, 2019; Kret et al., 2014). Based on Mathôt's (2013) approach, we reconstructed the pupil sizes for blinks (each window extended by 10 ms before and after) as well as gaps, using cubic spline interpolation. Beforehand, we filtered trials that had no pupil size data during the first or last 20 ms of the trial, so that sufficient data would be available for the baseline. The last 20 ms were chosen to keep the data symmetrical, and because the highest peaks in cognitive load would likely occur shortly before the task response, taking into account the delay in pupil dilation (see above, Mathôt et al., 2015). After that, the median pupil size during the first 40 ms was calculated as the baseline value for each trial (see Mathôt, 2013). A 10-point moving average filter was run to smooth the data for noise, the maximum pupil size was determined for each trial, and the difference to the baseline was calculated. Toth and Campbell (2021) pointed out that using the value of the maximum pupil diameter per trial in the analysis with baseline correction is the "most robust eye-tracking metric of cognitive effort (…) as it does not fall victim to the fact that viewing times across stimuli were not controlled" (Toth & Campbell, 2021, p. 2; also see Aboyoun et al., 1998; Campbell et al., 2018). As Beatty and Lucero-Wagoner (2000) stated, most of the task-induced pupil size changes are below 0.5 mm. To also account for rarer cases, we excluded differences larger than 0.6 mm. After that, the pupil size change was averaged between the two eyes. In case of available data for only one eye, this value was taken, due to the diameters of both eyes being highly correlated, especially locally (Jackson & Sirois, 2009). Overall, out of 10080 trials, 233 had missing data from both eyes and were excluded from analysis. In line with Campbell et al. (2018), we analyzed the maximum pupil diameter changes according to common behavioral data analysis, that is, excluding mirrored items and wrongly answered ones.

4.3.8 Statistical Analysis

The first part of this section provides a short overview regarding the statistical analysis approach for the three studies of this thesis. During the last decade, linear mixed models and general mixed models gained popularity due to increasing computing power and available packages for them (Bates, Maechler, et al., 2015). A lot of literature lay out the general advantages of mixed models, however, there are still aspects of the analysis that are being discussed, such as the choice of random slopes in the models (Barr et al., 2013; Bates, Kliegl, et al., 2015; Brauer & Curtin, 2018; Matuschek et al., 2017; see Jost, 2022). It is important to note that these analysis types are still evolving, and that there is an ongoing discussion about model selection depending on theory or data, and whether simple or complex models should be chosen (Jost et al., 2023). Thus, although applying state-of-the-art statistical analysis in the studies presented here, a possible uncertainty of the outcomes has to be taken into account (see Jost, 2022).

The general idea of mixed model analysis will be outlined and described: Mixed models incorporate fixed effects and random effects. Neither of them are clearly defined, but typically, fixed effects are the ones that describe the variables of interest (e.g., all treatments of interest in an experiment) and random effects are described to generalize the model to a more general sample (e.g., generalizing data of a test sample to a broader population). The random effects can also incorporate random intercepts that account for baseline differences in the data, and random slopes that account for possible differences of how the fixed and random effects interact. Mixed models offer various advantages, for instance, they allow the analysis of multiple crossed or nested random effects, the analysis of data that has unbalanced or missing values, the analysis of multiple covariates, and gaining higher statistical power. Additionally, with the simultaneous analysis of by-participant and by-item variances, averaging over participants or items is not necessary (Barr et al., 2013; Hilbert et al., 2019; Jost, 2022). Furthermore, gradual changes over the course of an experiment can be included (Baayen et al., 2008; Mirman et al., 2008; Winter & Wieling, 2016).

The following part provides details for the analysis of the first study. Statistical analysis was performed using *lme4* package (Bates, Maechler, et al., 2015) in R (R Core Team, 2018). Reaction time and pupil diameter were analyzed using linear mixed models and accuracy was analyzed using generalized linear mixed models with a binomial distribution. Model parameters were estimated by maximum likelihood estimation. *P* values were obtained by using likelihood ratio tests to test for improvement of model fit by the fixed effect of interest and compared to a significance level of .05. To date, no consensus exists on computing standardized effect sizes in linear mixed models, as they are generally used and reported for power- and meta-analyses (Feingold, 2009; Hedges, 2007; Jost, 2022; Rights & Sterba, 2019). For the significant effects and main effects, we report the unstandardized effect sizes and the confidence intervals that were calculated using parametric bootstrapping with 1000 simulations, in line with recommendations of Baguley (2009), and Pek and Flora (2018). Visual inspection of residual plots did not reveal deviations from homoscedasticity or normality in any model.

Hypothesis-driven model building was based on the research of Barr et al. (2013), and Bates, Kliegl, et al. (2015), starting with a model with random intercepts and slopes for every appropriate fixed effect and reducing the model complexity by dropping non-significant variance components. Non-significant fixed effects were further removed from the model, such that non-significant effects were tested for an improvement of model fit by inclusion in the resulting model while significant effects were tested for worsening of model fit by exclusion of the effect. Main effects for significant interactions were tested separately by splitting the interaction (see also Jost & Jansen, 2020). The resulting models for each parameter are described in the results section. All data were visualized using *ggplot2* package (Wickham, 2016) in R (R Core Team, 2018).

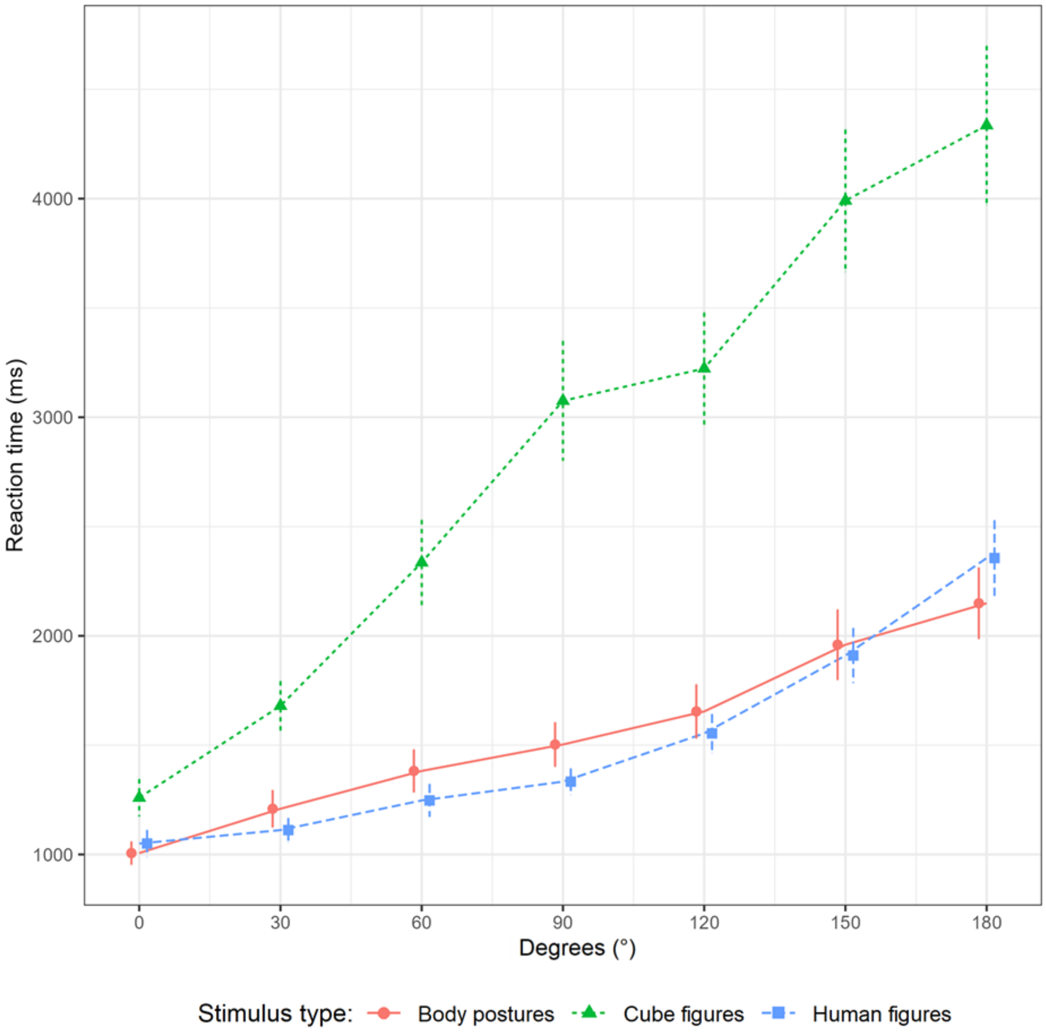
4.4 Results

4.4.1 Reaction Time

As shown in Figure 3, the reaction times for cube figures are constantly higher than for body postures and human figures. The graphs of the two embodied stimulus types are more closely related. All graphs show a positive slope with increasing angular disparity.

Figure 3

Reaction Time Plotted Against Angular Disparity and Stimulus Type.



Note. Dots represent the mean. Bars represent the standard deviation. Both are slightly offset for the groups for better visibility.

Model construction resulted in a model with random intercepts and slopes for STI and DEG by participant. STI*SEX*DEG and all respective interactions and main effects were analyzed as fixed effects. Significant differences were found for STI*DEG (see Table 1).

Reaction time increased significantly by DEG and STI (main effects). The interaction DEG*STI showed a significant increase in reaction time with increasing DEG for all stimulus types, with CF having the highest increase, followed by HF and BP. Pairwise comparisons for the interactions showed significant differences between all of them. Pairwise comparisons for the main effects showed significant differences between CF and BP, and CF and HF, but not for HF and BP, with CF always having higher values.

Regarding reaction time and our first hypothesis, sex differences emerged neither for the main effect (SEX) nor for the interactions. The results support our second hypothesis that mental rotation task performance would be better with embodied figures compared to the cube figures, which is shown both by the overall difference between the reaction times of the two embodied figure groups and the abstract figure one (STI; highest for CF) as well as the increase of reaction time for each degree of angular disparity (DEG*STI; highest for CF).

Table 1

Statistical Analysis of Reaction Time (in Seconds).

Variable	Estimate	SE	Test Statistic	*p* value	95% CI
Intercept	1.00	0.05			0.89, 1.10
DEG*STI			$\chi^2(2)=522.33$	<.001	
DEG*STI(BP)	0.59	0.05			0.49, 0.68
DEG*STI(HF-BP)	0.09	0.04	$\chi^2(1)=7.29$	<.001	0.01, 0.18
DEG*STI(CF-BP)	1.02	0.05	$\chi^2(1)=367.87$	<.001	0.93, 1.11
DEG*STI(CF-HF)	0.93	0.05	$\chi^2(1)=301.95$	<.001	0.83, 1.03
DEG(0°)*STI(HF-BP)	-0.12	0.06			-0.23, -0.01
DEG(0°)*STI(CF-BP)	0.19	0.08			0.03, 0.35
Main Effects					
DEG (100°)	0.91	0.04	$\chi^2(1)=151.84$	<.001	0.83, 1.00
STI			$\chi^2(2)=105.33$	<.001	
STI(HF-BP)	-0.03	0.04	$\chi^2(1)=0.89$	.345	-0.11, 0.04
STI(CF-BP)	1.03	0.07	$\chi^2(1)=103.22$	<.001	0.90, 1.17
STI(CF-HF)	1.07	0.08	$\chi^2(1)=100.42$	<.001	0.03, 1.21
Non-significant Effects					
SEX (male-female)	0.02	0.06	$\chi^2(1)=0.09$	.762	-0.10, 0.14
SEX*DEG	0.09	0.07	$\chi^2(1)=1.47$	.225	-0.06, 0.22
SEX*STI			$\chi^2(2)=2.05$	.360	
SEX*STI*DEG			$\chi^2(2)=4.11$	.128	

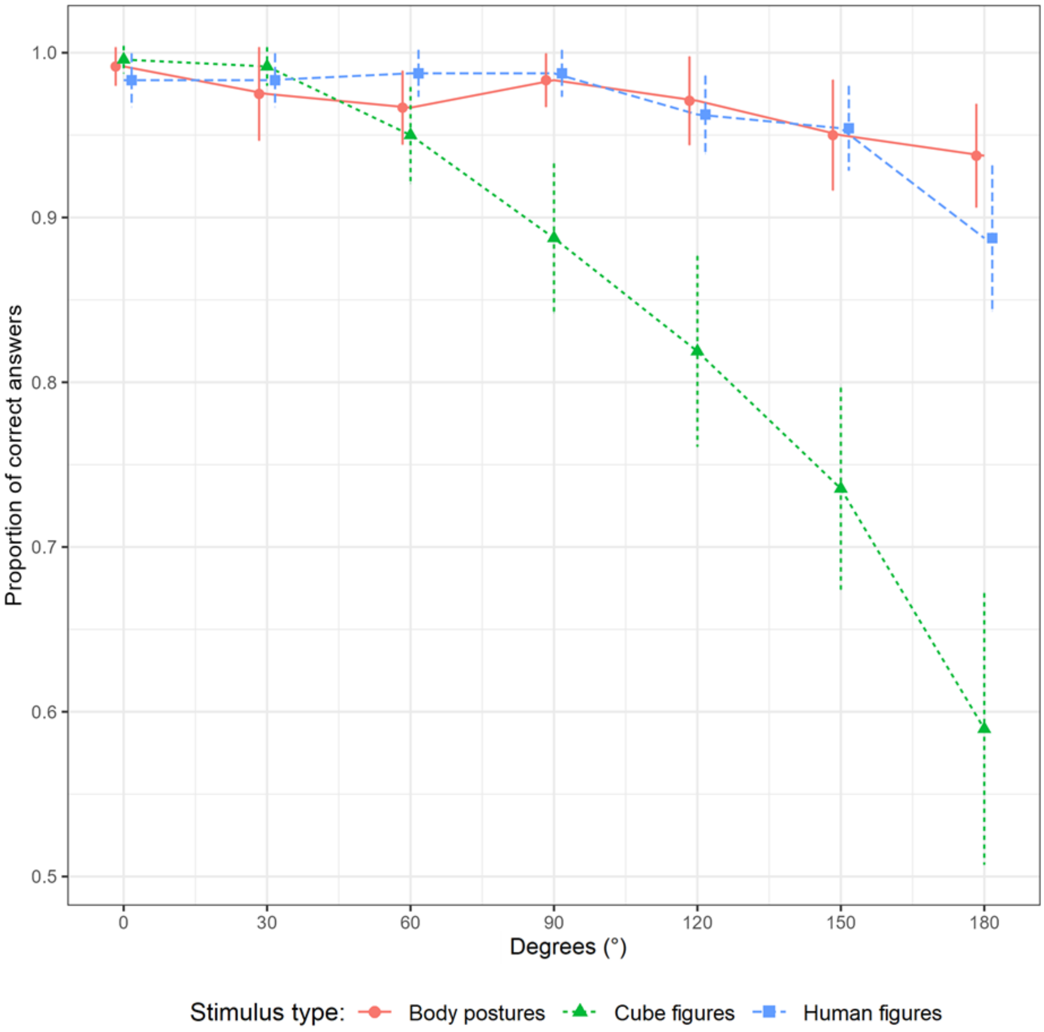
Note. Intercept in this model represents the estimate at 0° for body postures (BP). Effects of angular disparity (DEG) represent changes of 100°. Test statistic and *p* value for stimulus type (STI) and DEG*STI represent the pairwise comparisons (HF: human figures; CF: cube figures).

4.4.2 Accuracy

In Figure 4, the proportion of correct answers shows a steeper decline for cube figures than for body postures and human figures. The graphs of the two embodied stimulus types are more closely related. All graphs start similarly at zero degrees and show a negative slope with increasing angular disparity.

Figure 4

Accuracy Plotted Against Angular Disparity and Stimulus Type.



Note. Dots represent the mean. Bars represent the standard deviation. Both are slightly offset for the groups for better visibility.

Model construction for ACC resulted in a model with random intercepts and slopes for STI and DEG by participant. STI*SEX*DEG and all respective interactions and main effects were analyzed as fixed effects. Significant differences were found for STI*DEG (see Table 2).

Accuracy decreased significantly by DEG and STI (main effects). The interaction DEG*STI showed a significant decrease in accuracy with increasing DEG for all stimuli, with CF having the highest decrease, followed by HF and BP. Pairwise comparisons for the interactions showed significant differences between CF and BP, and CF and HF (with CF having larger decreases), but not for HF and BP. Pairwise comparisons for the main effects showed significant differences between all of them (with CF having lower values).

With regard to accuracy and our first hypothesis, sex differences emerged neither for the main effect (SEX) nor for the interactions. The results also support our second hypothesis that task performance would be better with embodied figures, which is shown both by the overall difference between the accuracy of the two embodied figure groups and the abstract figure one (STI; lowest for CF) as well as the decrease of accuracy for each degree of angular disparity (DEG*STI; largest for CF). The results are similar to those for reactions times, except for the pairwise comparison of the main effect (STI(HF-BP); significant differences in accuracy, but not in reaction time) and the slope by DEG (DEG*STI(HF-BP); significant differences in reaction time, but not in accuracy) between the two embodied figure groups.

Table 2

Statistical Analysis of (Logarithmic Odds of) Accuracy.

Variable	Estimate	SE	Test Statistic	*p* value	95% CI
Intercept	5.47	0.58			4.54, 7.07
DEG*STI			$\chi^2(2)=14.36$	<.001	
DEG*STI(BP)	-1.04	0.35			-1.85, -0.36
DEG*STI(HF-BP)	-0.53	0.45	$\chi^2(1)=2.87$	.090	-1.46, 0.35
DEG*STI(CF-BP)	-1.36	0.40	$\chi^2(1)=9.65$	.002	-2.23, -0.55
DEG*STI(CF-HF)	-1.03	0.41	$\chi^2(1)=6.33$	.012	-1.84, -0.24
DEG(0°)*STI(HF-BP)	-0.15	0.71			-1.67, 1.18
DEG(0°)*STI(CF-BP)	-0.67	0.64			-2.23, 0.42
Main Effects					
DEG (100°)	-1.95	0.18	$\chi^2(1)=94.15$	<.001	-2.33, -1.61
STI			$\chi^2(2)=66.44$	<.001	
STI(HF-BP)	-0.93	0.41	$\chi^2(1)=5.24$	.022	-2.15, -0.21
STI(CF-BP)	-2.58	0.39	$\chi^2(1)=64.54$	<.001	-3.72, -1.96
STI(CF-HF)	-1.69	0.25	$\chi^2(1)=46.50$	<.001	-2.31, -1.23
Non-significant Effects					
SEX (male-female)	0.24	0.21	$\chi^2(1)=1.29$	.256	-0.18, 0.59
SEX*DEG	-0.05	0.31	$\chi^2(1)=0.03$	.863	-0.68, 0.56
SEX*STI			$\chi^2(2)=2.22$	.330	
SEX*STI*DEG			$\chi^2(2)=1.21$	.545	

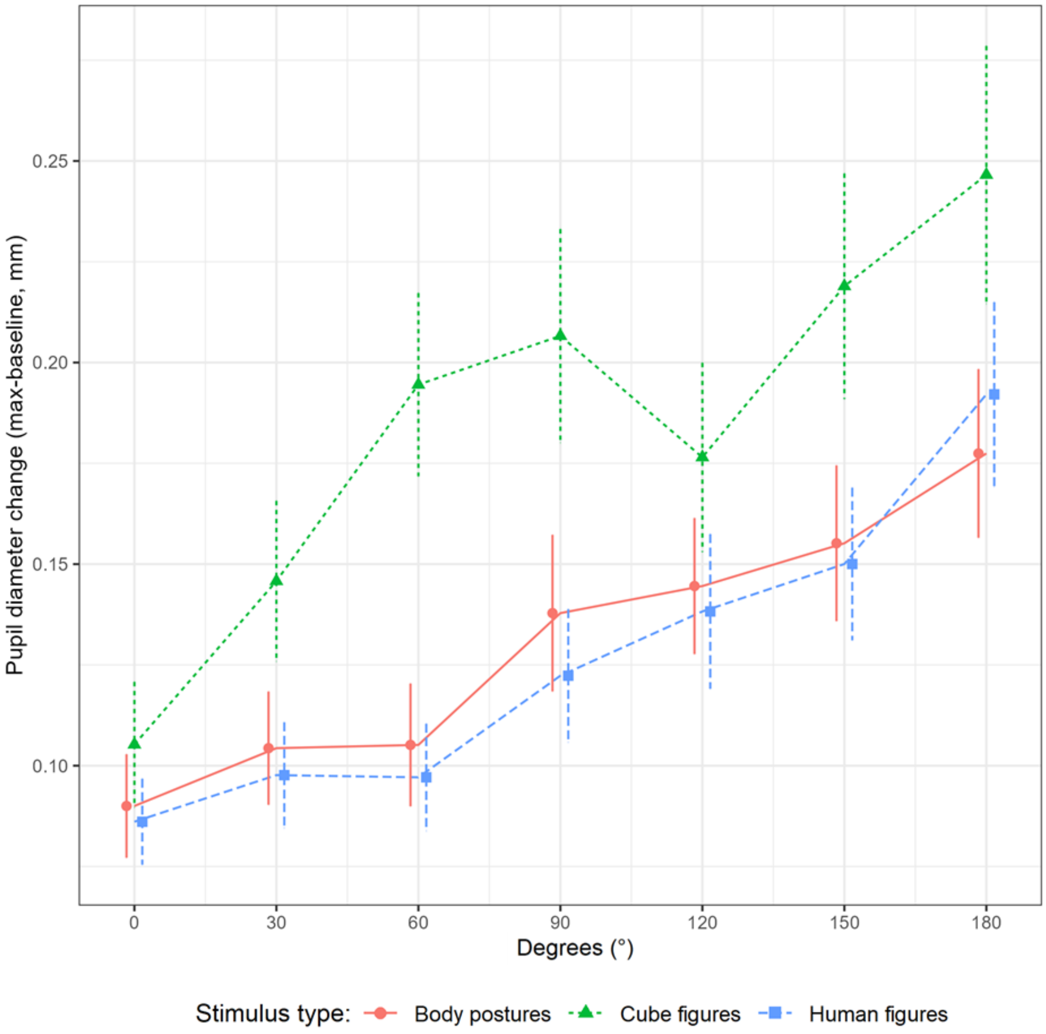
Note. Intercept in this model represents the estimate for the logarithmic odds at 0° for body postures (BP). Effects of angular disparity (DEG) represent changes of 100°. Test statistic and *p* value for stimulus type (STI) and DEG*STI represent the pairwise comparisons (HF: human figures; CF: cube figures).

4.4.3 Pupil Diameter

The task-evoked pupil responses are depicted in Figure 5 and show an inclination of all graphs with increasing angular disparity. Here, cube figures show constantly higher values than the two embodied figure types. The graphs of body postures and human figures lie closer together.

Figure 5

Changes of Pupil Size (Max-Baseline) Plotted Against Angular Disparity and Stimulus Type.



Note. Dots represent the mean. Bars represent the standard deviation. Both are slightly offset for the groups for better visibility.

The model building resulted in a model with random intercepts and random slopes for DEG and RT by participant. STI*SEX*DEG, RT, and all respective interactions and main effects were analyzed as fixed effects. The pupil diameter increased significantly by STI*DEG, and RT (see Table 3). Pupil size increased significantly by DEG and STI (main effects). The interaction STI*DEG showed a significant increase in pupil size for each DEG for the different stimuli, with HF having the highest increase, followed by BP and CF. Pairwise comparisons for the interactions showed significant differences between CF and HF, but not for CF and BP, and HF and BP. Pairwise comparisons for the main effects showed significant differences between CF and BP, and CF and HF, but not for HF and BP. With regard to the inclusion of RT in the model, we inspected the data for possible collinearity problems. All variance inflation factors were smaller than three (maximum of 1.86), that means, collinearity was not an issue (see Zuur et al., 2010).

In the results for changes of pupil size, there was evidence for our third hypothesis that the pupil dilation would be influenced by the stimulus type and angular disparity. This is shown both by the increase of the pupil size change with increasing angular disparity (DEG*STI) as well as the overall difference between the pupil size change of the two embodied figure groups and the abstract figure one (STI). Here, the abstract figures have the lowest slope for DEG*STI, which could partially result from their constantly higher values in pupil size change (STI). There was no evidence for our fourth hypothesis, as no sex differences emerged neither for the main effect (SEX) nor for the interactions.

Table 3

Statistical Analysis of Pupil Diameter (in 10^{-1} mm).

Variable	Estimate	SE	Test Statistic	*p* value	95% CI
Intercept	1.41	0.06			1.27, 1.53
DEG*STI			$\chi^2(2)=8.72$	.013	
DEG*STI(BP)	0.28	0.05			0.19, 0.37
DEG*STI(HF-BP)	0.06	0.06	$\chi^2(1)=0.71$	.399	0.06, 0.17
DEG*STI(CF-BP)	-0.13	0.06	$\chi^2(1)=1.33$	.250	-0.25, 0.00
DEG*STI(CF-HF)	-0.16	0.07	$\chi^2(1)=5.98$	.015	-0.30, -0.03
Main Effects					
RT (sec)	0.37	0.03	$\chi^2(1)=97.59$	<.001	0.31, 0.43
DEG (100°)	0.28	0.04	$\chi^2(1)=45.74$	<.001	0.21, 0.35
STI			$\chi^2(2)=40.17$	<.001	
STI(HF-BP)	-0.03	0.03	$\chi^2(1)=0.64$	.423	-0.09, 0.04
STI(CF-BP)	0.22	0.04	$\chi^2(1)=35.09$	<.001	0.14, 0.30
STI(CF-HF)	0.27	0.04	$\chi^2(1)=39.18$	<.001	0.18, 0.36
Non-significant Effects					
SEX (male-female)	-0.09	0.08	$\chi^2(1)=0.10$	.752	-0.25, 0.07
SEX*DEG	-0.09	0.07	$\chi^2(1)=2.82$	.093	-0.23, 0.05
SEX*STI			$\chi^2(2)=5.47$	.065	
SEX*STI*DEG			$\chi^2(2)=0.15$	.928	

Note. Intercept in this model represents the estimate for body postures (BP), average reaction time (RT), and average angular disparity (DEG). Effects of DEG represent changes of 100°. Test statistic and *p* value for stimulus type (STI) represent the pairwise comparisons (HF: human figures; CF: cube figures).

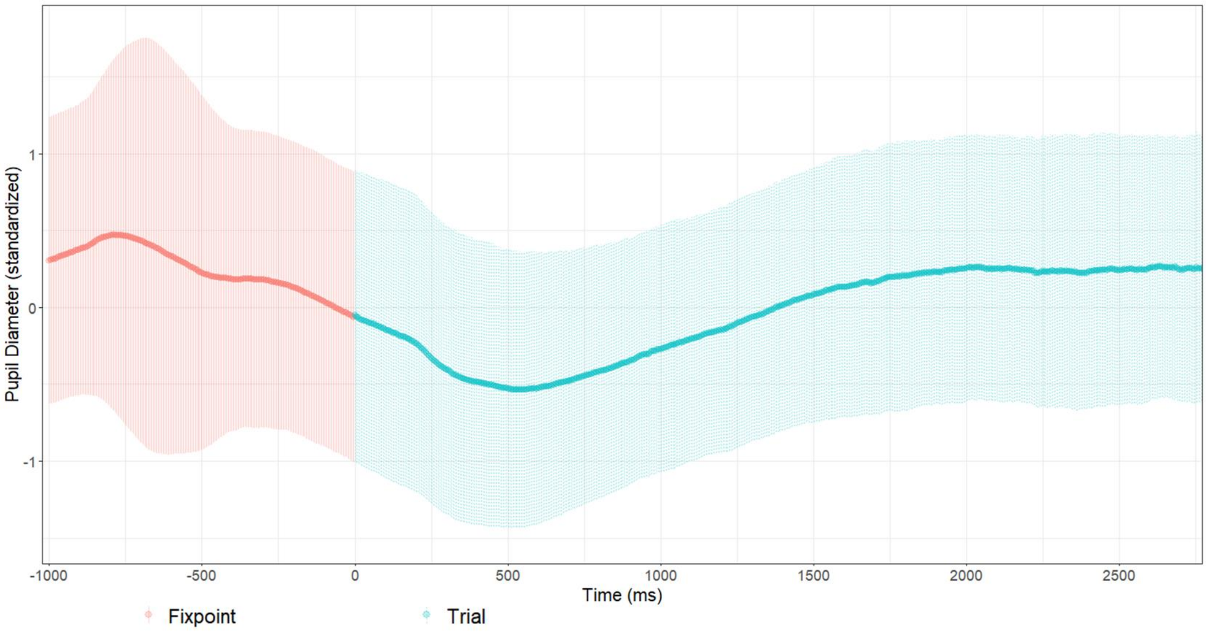
4.4.4 Exploratory Results

We exploratively analyzed the time course of pupil dilation throughout the trials. Besides gaining insight into time-dependent pupil size changes of interest for future studies, this can also help to check the validity of the data, both as a check for the selected pupil size change measures (baseline and maximum) as well as possible disturbances by, for example, carry-over effects.

The time course of pupil dilation (see Figure 6) illustrated that the pupil size decreased after trial start for around 500 ms and began to increase thereafter. Alignment of the trial ends illustrated that the pupil diameter increased during the time around task response, leading to higher values during the fixation point (see Figure 7).

Figure 6

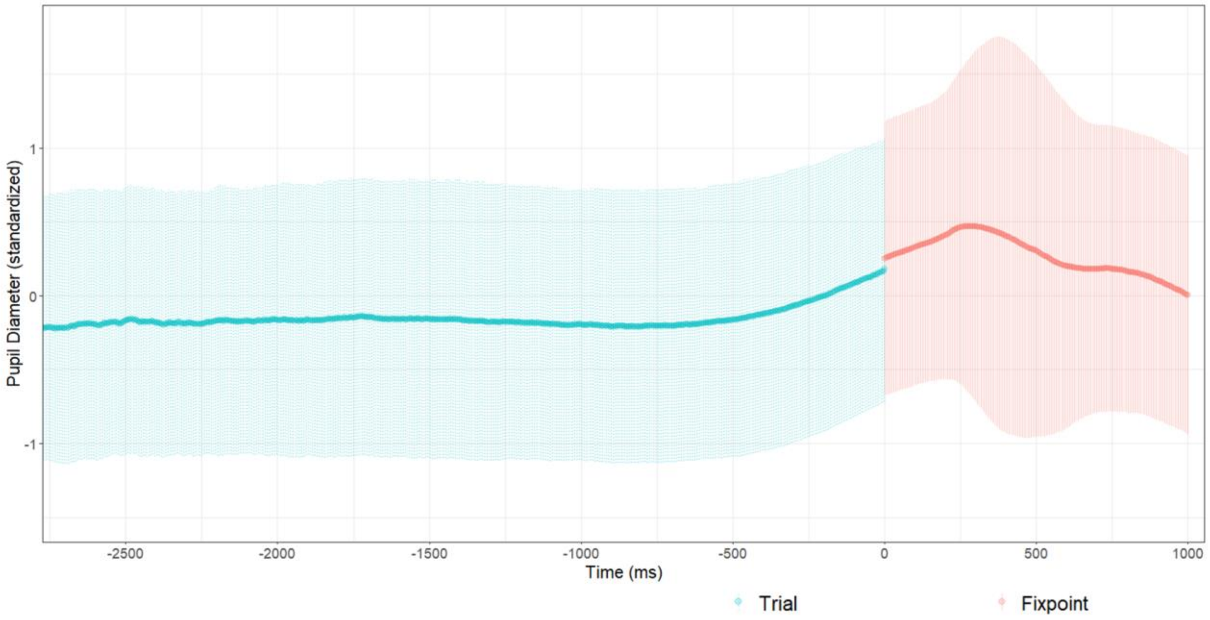
Time Course of the Standardized Pupil Diameter Means of All Participants' Trials, Depicting Only the First Three Seconds.



Note. Data of shorter trials only go until their respective trial end. Zero represents the onset of the stimulus preceded by the fixation point (−1000 ms to 0 ms). The shaded areas represent plus/minus one standard deviation.

Figure 7

Time Course of the Standardized Pupil Diameter Means of All Participants' Trials, Depicting Only the Last Three Seconds.



Note. Data are aligned for trial ends at 0 ms; data of shorter trials only begin at their respective trial start. Zero represents the end of the stimuli and is followed by the fixation point (0 ms to 1000 ms). The shaded areas represent plus/minus one standard deviation.

4.5 Discussion

4.5.1 Validity of Pupil Size Measurements

Before addressing the hypotheses, we would like to elaborate on the temporal analysis of the pupil size means. Visual inspection of the time course of pupil dilation indicated possible carry-over effects. Over all trials, pupil size decreased after trial start for 500 ms and began to increase thereafter. Therefore, the one second fixation point duration was seemingly not enough for the pupil diameter to fully return to its baseline value. Inspecting the end of trials indicated that the pupils were still dilating before, while, and after task response (mouse-click) was given, leading to higher values during the fixation point. This is a possible concern to the validity of the measurements as the exact extent and duration of the observed effects are unknown. One possible explanation is that shorter trials were systematically affected more, as the recording was cut off earlier and the trailing pupil dilation would lead into the following fixation point and trial. On longer trials, there was more time for the pupil size to plateau, which could be a reason for our results indicating that higher values emerged for higher angular disparity. This explanation however is unlikely, as this should have been compensated by the included effect of reaction time on pupil dilation. Nevertheless, assuming that carry-over effects only depend on the difficulty of the previous trial (and trials are in random order) and are somewhat random in magnitude and duration, they at least introduce additional variance to the measurements, which in turn reduces the power of the design.

Despite these issues, we did not choose to alter the measurements. Regarding the baseline measurement, we cannot isolate a baseline due to the possible overlap in time of the observed decrease in pupil size and the expected increase due to the cognitive effort. However, we also conducted the pupil diameter analysis using baseline values for each stimulus, which were recorded at the beginning of the experiment. This form of analysis has its own problems, that is, mainly having no control of the random trial by trial pupil fluctuations. Interestingly, this and our described analysis results did not differ in the significance of any effect in question. This could indicate that the starting values were overall only shifted, that is, baseline measurements at the start of each trial were always 0.5 units larger than the true baseline. Regarding the measurement of maximal pupil dilation, one could include the following fixation point in the analysis of each trial. This however would have other effects influencing the data, for example, new visual input and pupillary light reflex. In addition, altering the analysis would not change the problem of the carry-over effects. In consequence, regarding the hypotheses for the pupillometric measurements, the results must be considered with caution, independent of a possible change of measurements. Nevertheless, these issues are important for both past and future studies of pupil dilation during mental rotation and also other cognitive tasks where similar problems might arise. In experiments where the trial was cut off directly after task response, the interpretation of the results has to be done cautiously, as carry-over effects might have a similar impact there (e.g. in, Campbell et al., 2018). In a recent study using another approach, Bochynska et al. (2021) showed the stimulus for four seconds, independent of task response. However, since trials with response times longer than four seconds (141 of 1064 trials) were excluded from the analysis, tasks of higher angular disparity might only have been partially included, and the ones included could also have faced the problems of delayed pupil dilation.

Based on the observed time course of pupil dilation, future studies should 1) keep showing the task and measuring pupil dilation even after the response for at least 500 ms, and 2) increase the break between trials to at least two seconds.

4.5.2 General Discussion

The results show no sex differences in the behavioral performance. The main effects and interactions for both accuracy and reaction time do not show any influence of sex on the models. This is in line with other chronometric mental rotation studies. For instance, Voyer et al. (2006) also report no sex differences in mental rotation performance of 3D cube figures. Jansen-Osmann and Heil (2007) investigated sex differences in mental rotation tasks with five different stimulus types and also reported sex differences only in one (polygons) of these (3D cube figures, letters, stimuli from primary mental abilities, and animal pictures). However, the results are in contrast to the study of Voyer and Jansen (2016)—using stimuli that were partially the same as in this study—who pointed out that although a performance improvement for both sexes was apparent, men might benefit more from the advantage through embodiment. One reason for this discrepancy may be the variation in the use of the human stimuli between the two studies. Voyer and Jansen (2016) presented head cubes (cubes with the addition of a head) while we investigated human postures.

In accordance with Amorim et al. (2006), our results confirm our second hypothesis that task performance would be better for both embodied figures compared to cube figures for both sexes. Our experiment shows a significant embodiment effect. Both embodied figures in this object-based transformation task were processed more easily on a behavioral level (shorter reaction time and higher accuracy), matching our hypothesis. This is in congruence with other studies (e.g., Amorim et al., 2006; Campbell et al., 2018; Voyer & Jansen, 2016).

In line with the paradigm for chronometric mental rotation tasks, changes in angular disparity significantly influenced all dependent variables for all models. Higher angular disparity between the two pictures resulted in higher reaction times and lower accuracy. These effects are larger for the abstract than for the embodied figures. In particular, the interaction of cube figures and angular disparity showed a higher negative impact on performance than for body postures and human figures.

In terms of cognitive load (Hypothesis 3), cube figures show the highest values in pupil diameter, followed by body postures and human figures. Here, both embodied figure types differ significantly from the cube figures. Therefore, the highest cognitive load manifests in cube figures, indicating that these tasks are more difficult to solve. This finding is congruent with our results for behavioral performance.

An additional possible explanation for the pupil diameter to be lower for both embodied figures than for cube figures might be a congruency effect. In a pupillometry experiment regarding the Stroop task, Hershman and Henik (2019) report lower pupil diameter values for neutral (colored letters with no meaning) than for color-congruent (word and word color align) tasks, indicating higher cognitive load in the latter due to a task conflict between reading the word and naming the color of the word. This is based on the effect that stimuli evoke tasks, which are strongly associated with them (Rogers & Monsell, 1995; Waszak et al., 2003). That means the neutral colored word has more task congruency, because it only elicits naming the color, whereas the colored word elicits reading of the word, creating conflict with the response. These findings concur with the description of motoric embodiment (i.e., imagination and execution of actions addressing the same motor representations) of Amorim et al. (2006) and might also apply to this study. Task response was given in interaction with a desktop mouse. With the hand being an important and salient feature of human figures and body postures, a congruency effect with the hand response is possible. That is, higher congruency could also lead to lower cognitive load and might thus be an additional factor considering motoric embodiment effects in this experiment.

Overall, the first part of our hypothesis 3 was confirmed by the modeling results, with the main effects of cube figures and higher angular disparity increasing the pupil dilation the most. Also as predicted, the two embodied figure types did not differ significantly. Interestingly, the interaction of stimulus type and angular disparity showed the highest increases for human figures. With the implementation of reaction time in the model, the results illustrate the additional effect of this interaction on top of the effect of reaction time. Here, the pupil sizes for cube figures are higher due to the higher difficulty to solve the task. Thus, the range to increase was smaller than for the embodied figures, resulting in a less steep slope by angular disparity, which might indicate a ceiling effect in this regard. However, no significant differences emerged between body postures, and cube and human figures, with the values of body postures lying between those two. Consequently, they cannot be placed properly in this regard, which should be further investigated in future experiments. Additionally, changes in pupil dilation did not get fully explained by variations in angular disparity. Reaction time itself still predicted a significant portion of the pupil diameter in the statistic model. That is, both reaction time and angular disparity had an impact on cognitive load, but none of them alone seemed to be sufficient to describe the connection between each other, and to account for task difficulty. As a consequence, it is possible that the relationship between difficulty and cognitive load is not linear.

Our expectations regarding the sex differences in pupil dilation were based on the findings of Campbell et al. (2018), but our results did not confirm these predictions, as no differences emerge. As a conclusion, both sexes showed similar performance in the mental rotation tasks and exerted the same levels of cognitive effort for each stimulus type. This also means that all embodiment facilitations provoked similar changes in cognitive load and task performance for both sexes. This gives a hint that both men and women are able to solve mental rotations tasks, which have the same amount of spatial embodiment, with the same effort (Amorim et al., 2006). Spatial embodiment includes the mapping of a body-relevant coordinate system that facilitates the mental rotation process. In our case, having controlled object-based transformations as instructed, spatial embodiment seems to partially explain the differences in behavioral performance. With the results for pupil dilation matching those for behavioral data, spatial embodiment can explain the reduced pupil size for embodied figures, having a logical alignment in the geometric space.

Our findings did not show any differences between the two embodied stimulus types. This is interesting, as they differ between each other regarding their geometric form. Body postures and cube figures are s-shaped comprising three bends, with body postures imitating the shape of the abstract figures. Human figures instead comprise only one or two bends, and are more i- or t-shaped. Thus, the stimulus types differ in spatial complexity, and human figures could theoretically be assumed to facilitate visual absorption and processing. However, the difference in geometric form had no impact in this experiment. Overall, our results provide—for the first time—evidence that men and women need the same cognitive effort for the solution of this task.

4.6 Limitations

In pupillometry studies, the pupil foreshortening error should be of concern to have an influence on the acquired data, especially when looking at areas, which are further away from the screen center (Hayes & Petrov, 2016). In our mental rotation tasks, all analyzed image pairs are shown at the same positions on the screen. Possible pupil foreshortening errors were not expected to systematically differ between conditions, and were neglectable for that reason (see Mathôt et al., 2018).

Amorim et al. (2006) neither analyzed sex differences nor elaborated on their choice of color. The partially pink coloring of the stimuli might facilitate female performance due to possible familiarity effects. However, a similar effect could also apply to males in respect to cube figures (Ruthsatz et al., 2017). To further analyze effects of stimulus color, the same stimuli would have to be presented in different color schemes (e.g., black-white, blue, and pink), and analyzed regarding sex differences.

In this experiment, no chinrest was used. Although the SMI software is able to compensate for variations in head-tracker distance, these estimations might have limitations. Those can result in random noise, which does not endanger the validity of the results, but could be avoided nonetheless. Thus, in pupillometry experiments, using a chinrest is recommendable.

As mentioned in the discussion, the experimental design led to possible carry-over effects, which have to be taken into account for interpretation of the results.

4.7 Conclusion

This study replicated the effect for mental rotation tasks of embodied stimuli to be easier to process (Amorim et al., 2006) and showed for the first time that this result is confirmed with pupillometry, meaning that mental rotation of embodied figures needs less cognitive effort to solve the task. Sex differences did not appear in any of the measurements.

5 Second Study: The Effect of a Mindfulness Session on Mental Rotation Task Performance and Subjective Cognitive Load

5.1 Summary

In the second study, we investigated the effects of a single bout of a brief (20-minute) body-scan intervention on chronometric mental rotation task performance and subjective cognitive load. 91 participants (44 women and 47 men) were divided into two test groups (mindful or control intervention). They completed two sets of 150 items each with cube figures of similar complexity and filled out a test of state mindfulness. Subjective cognitive load was measured after each item block with a Likert-type scale. Subjective cognitive load, reaction time, and accuracy were analyzed with linear mixed models with the factors time, mindfulness, angular disparity, and sex. The mindfulness induction led to significantly elevated state mindfulness in comparison to the control condition. The main effects of the mindfulness induction were not significant in any of the dependent variables; however, a significant three-way interaction emerged between pre-post testing, mindfulness, and sex for reaction times. This interaction indicated that women might have benefited more from the mindfulness induction and men more from the control condition. Furthermore, men's performance was more affected by increasing levels of angular disparity than women's. Analysis of subjective measures of cognitive load showed reduced load in the posttest as well as a connection to reaction times. In the model, larger sums of reaction times would lead to significantly higher levels of subjective cognitive load.

5.2 Goals and Hypotheses

With mathematical abilities and mental rotation performance being related (Xie et al., 2020) and mental rotation also being influenced by embodiment effects through mindfulness induction (e.g., Rahe & Jansen, 2023), the reported effects of mindfulness in mathematics (Weger et al., 2012) might also apply to mental rotation tasks. Thus, our goal was to conduct an online study to investigate the effects of an embodied mindfulness treatment on chronometric mental rotation. Participants of both genders performed mental rotation tasks and were assigned to two condition groups, with or without mindfulness intervention. Additionally, we used a scale for perceived cognitive load to investigate if the sex differences in mental rotation are connected to subjective cognitive load. The following hypotheses were investigated:

In line with former research indicating task performance improvements (e.g., Weger et al., 2012), we predicted an interaction between pre-post testing and mindfulness, with mindfulness improving the mental rotation performance in the posttest compared to the control group (Hypothesis 1). We expected these changes to be larger for tasks of higher difficulty, that is, a higher angular disparity between both objects being compared (see Jost & Jansen, 2020).

With regard to the benefits of focused attention meditation (see Lutz et al., 2008), mindfulness was expected to lead to lower subjective cognitive load in the posttest (Hypothesis 2).

Apart from the treatment effects, we expected a general decline in subjective cognitive load from pre to posttest due to possible learning and habituation effects (Hypothesis 3).

According to the first study of this thesis (Bauer et al., 2022), we expected subjective cognitive load for pre- and posttests to be linked to their overall sum of reaction times. Longer overall response times were expected to lead to higher values in cognitive load (Hypothesis 4).

In form of additional hypotheses, we expected a time effect of angular disparity, that is, higher behavioral performance improvements in the posttest for higher angular disparity. According to the paradigm of mental rotation, we expected reaction times to increase and accuracy rates to decrease with higher angular disparity. The effect for sex differences will be analyzed, too, exploratively. Additionally, we also analyzed exploratively possible sex differences in overall reaction time sums of the test blocks.

5.3 Methods

5.3.1 Participants

In the study, 149 students (84 women) participated and received study credits. In the preprocessing of the data, various filters led the data of 58 participants to be excluded from further analysis (for details, see ***5.3.9 Data Processing***). Consequently, 91 students (44 women, mean age (SD) = 21.8 (1.2) years; 47 men, mean age (SD) = 22.8 (3.5) years) form the sample for statistical analysis. In the study of Weger et al. (2012), the main effect of mindfulness on mental rotation for women was detected with a $d = .58$. The required sample size for women was estimated using the software G*Power (Faul et al., 2007). With an effect size of $d = .58$ ($f = .29$), a power of 0.90, and an alpha value of .05, a sample size of 66 women was estimated (see Brysbaert & Stevens, 2018). For the explorative analysis of sex differences, the same amount was estimated for men. Due to the filtering, the resulting sample was therefore smaller than estimated *a priori*. However, the power should increase by using linear mixed models in the analysis (Barr et al., 2013; Hilbert et al., 2019).

Table 4

Participants' Experience in Mindfulness, Meditation, Yoga, and Mental Rotation Tests Prior to This Study.

Experience	Mindfulness	Meditation	Yoga	Mental Rotation
None	22	10	4	19
Tried out once		40	32	
Done a few times per year		28	26	
Done a few times per month		10	22	
Done a few times per week		3	7	
Minor theoretical knowledge	37			
Practical experience	32			
Participated in 1-3 experiments				49
Participated in >3 experiments				23

The participants' experience in mindfulness, meditation, yoga, and mental rotation tests prior to this study is shown in Table 4. To mitigate possible experience effects, participants were required not to have participated in any mental rotation task experiment within the six months preceding this experiment. All participants reported no relevant physical or mental limitations. Informed consent was obtained from all individual participants. The experiment was conducted according to the ethical declaration of Helsinki. We communicated all considerations necessary to assess the question of ethical legitimacy of the study.

5.3.2 Conditions

Participants were randomly assigned to one of two test groups. The intervention after completing the first block of mental rotation tasks depended on this assigned group. In line with Ussher et al. (2014; ten-minute reading) and as in the studies of Aaron et al. (2020) and Schroter et al. (2023), the control group heard a twenty-minute reading about natural history in German ("A Short History of Nearly Everything"; Bryson, 2004). For the mindfulness condition, a twenty-minute audio track consisted of recorded instructions by a professional *Mindfulness-Based Stress Reduction* instructor with more than ten years of experience to do a mindfulness-based body-scan. The focus of the meditation was on the body as a whole, on body sensations, and on the perception of the hands. In the final sample with 91 participants, 25 women and men were assigned to the mindfulness condition. Nineteen women and 22 men were assigned to the control condition.

5.3.3 Setup

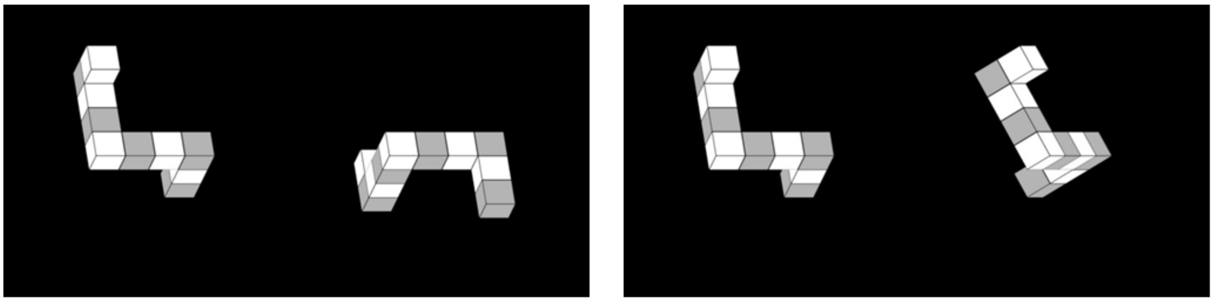
Stimulus presentation and response handling were controlled with OpenSesame software (version 3.3.1 *Lentiform Loewenfeld 2020*, Mathôt et al., 2012) using the OSWeb tool (version 1.3.8.0) together with JATOS software (version 3.5.4; Lange et al., 2015) and Ngrok software (version 2.3.35; ngrok Inc., 2023). With this setup, the study was run as an online experiment, which participants could access through their Internet browser. To mitigate possible confounding variables in the online experiment setup, various instructions were given to the participants beforehand. The participants were asked to prepare a room, where they could spend at least 90 minutes quietly and undisturbed. They were also advised to prepare a comfortable place, for instance, a bed, mattress, or sofa, where they could lie down during the twenty-minute session in the middle part of the experiment. They were also asked to put their phones on silent mode and make sure that their computer had a power supply connected as well as a sufficiently loud volume output and monitor brightness. All participants used keyboard arrow buttons for task response to control for input lag.

5.3.4 Stimuli

Stimuli consisted of a selection from the stimulus library of Peters and Battista (2008). Ten cube figure models with rotations around x- and z- axis in 45° steps and mirrored/non-mirrored orientation (i.e., figures one to ten with *a* and *b* orientation in the stimulus library) were used with a checkered pattern on black background (see Figure 8).

Figure 8

Examples of Mental Rotation Stimuli Used in the Experiment.



Note. Selected from the stimulus library of Peters and Battista (2008). Left: 90° rotation in x-axis. Right: 45° rotation in z-axis.

In the main experiment, which was divided in two parts, each part consisted of 150 stimuli, resulting in a total of 300 different stimuli. The first 150 stimuli were figures one to five, and the second 150 stimuli were figures six to ten of the stimulus library (Peters & Battista, 2008). The order of the stimuli in each block was randomized for all participants. On the left side of the screen, every model was presented in orientation *a*, rotated by 30° in x direction and 15° in z direction, so that the base model for x or z rotation was identical. On the right side of the screen, a rotated and mirrored/non-mirrored stimulus was presented. Stimulus pictures were sized 400px times 400px and presented vertically centered and horizontally positioned 300px to the left or right of the center of the screen until a response was given. Thus, depending on the monitor resolution, the stimuli were possibly slightly bigger or smaller for each participant. However, since the distance or position to the screen could not be controlled either, and as no pupillometric measurements are conducted in this setup, these differences can be regarded as negligible. In the practice block, three different models were presented in overall 24 practice stimuli (figures 11 to 13 of the stimulus library), that is, four rotations in x-axis and z-axis for each model, combined with using each possible angular disparity once and half of the trials being mirrored. In the practice session, between the stimuli pairs, participants received feedback for 1000 ms (✔- right, ✘- wrong) shown at the center of the screen. In experimental sessions, a fixation cross ("+") was shown there for 500 ms.

5.3.5 Subjective Cognitive Load

After the practice, pretest, and posttest block, participants filled out a Likert-type scale for subjective cognitive load. Here, they rated their response to each block on a ten-point scale. The measure of cognitive load was similar to Ayres (2006): "In the recent section, how demanding/ strenuous did you perceive solving the mental rotation tasks?" The answers were given as a number and a verbal expression, for example, "6 – Very demanding" (questions and answers were presented in German; see Table 5). Measuring the cognitive load for each block of tasks was chosen over the measurement for every single task. This was done to maintain the general mental rotation task design, without disruption of the consecutive tasks by other questions, and because the complete experiment duration would have been extended by about an hour for each participant. This is in line with Jost et al. (2023), who also used this type of scale to measure the RPE (rate of perceived exertion) for subjective cognitive exertion after each block in their study regarding mental rotation and aerobic exercise.

Table 5

German Scale of (Self-Reported) Subjective Cognitive Load Used in This Study (Left) and an English Translation (Right).

Wie anstrengend empfanden Sie es im eben durchgeführten Teil, die mentalen Rotationsaufgaben zu lösen?	In the recent section, how demanding/ strenuous did you perceive solving the mental rotation tasks?
9 - So anstrengend, dass ich sie nicht lösen konnte	9 - So demanding that I could not solve them
8 - Maximal anstrengend	8 - Maximally demanding
7 - Extrem anstrengend	7 - Extremely demanding
6 - Sehr anstrengend	6 - Very demanding
5 - Anstrengend	5 - Demanding
4 - Mäßig anstrengend	4 - Moderately demanding
3 - Wenig anstrengend	3 - Little demanding
2 - Sehr wenig anstrengend	2 - Very little demanding
1 - Extrem wenig anstrengend	1 - Extremely little demanding
0 - Überhaupt nicht anstrengend	0 - Not at all demanding

5.3.6 Toronto Mindfulness Scale

As a manipulation check whether the mindfulness manipulation had the desired effect, we used the Toronto Mindfulness Scale (TMS; Lau et al., 2006) as a measure of state mindfulness. The TMS was translated to German and included 13 items with five answer alternatives (0 = not at all; 4 = very much). For each participant, we computed summed scores for all items (see Lau et al., 2006). To prevent priming the participants to focus on mindfulness behavior prior to the treatment (either control or mindfulness condition), the TMS was only done after the second mental rotation task block. Thus, these values served as a check whether the treatments showed an overall different effect between the groups. In this study, the internal consistency of the questionnaire is good, with the Cronbach's alpha of the TMS score being 0.82.

5.3.7 Procedure

The experiment was a single session and lasted between 60 to 90 minutes, depending on participants' speed to complete all items. At the start of the online experiment, the participants were presented written information about the study protocol and the goals of the study. Then, they gave informed consent and were directed to a follow-up page with instructions for the following experiment. One of those was to set the browser window to full-screen presentation. After that, the practice session with feedback followed, which was introduced by a digitally presented instruction. Participants used the keyboard for response handling and received written instructions to press the left arrow button, if the stimuli could be rotated into congruence (non-mirrored, "same"), and the right arrow button, if the two stimuli were mirrored ("different"), and to answer as quickly and precisely as possible. After completing all practice trials, they answered the scale for cognitive load (see Table 5). Then, the first part of the main experiment was run, after which they answered the scale for cognitive load again. Depending on the assigned test groups, an intervention followed. This intervention was either a twenty-minute mindfulness intervention, during which the participants listened to recorded instructions to do a mindfulness-based body-scan, or a twenty-minute control intervention, during which the participants listened to a section of an audiobook about natural history. Then, the second part of the main experiment (posttest) was done. Following this part, they answered the scale for cognitive load, filled out the TMS (Lau et al., 2006), a demographic questionnaire, and the question whether they had honestly been attentive during the twenty-minute listening part (e.g., if they had fallen asleep or used their phones). Then, the participants received their participation code and the experiment concluded. All participations were conducted anonymously, and the presented participation codes in the end were distributed randomly from a predefined list of one thousand codes and were not recorded anywhere. As it was an online experiment, all instructions were standardized.

5.3.8 Study Design

To analyze cognitive performance, the dependent variables are reaction time (RT) and accuracy (ACC). Subjective cognitive load was self-reported on a Likert-type scale after the pre- and posttest. Independent variables for the analysis of the cognitive performance are pre-post experiment setup (PP), mindfulness intervention (MF), SEX, angular disparity (DEG), and their respective interactions. DEG describes the angular disparity between the two figures shown on the screen. DEG was included as a fixed effect as it is the main moderator of difficulty in mental rotation tasks (Jost & Jansen, 2020). Since the left image was always presented with 0° rotation, the angular disparity depicts the rotation in degrees of the right image. The cognitive load analysis includes PP, MF, SEX, and their respective interactions. Furthermore, the total sum of reaction time (for all items; RT-Sum) for each block (pre- and posttest) is analyzed as an additional fixed effect to analyze the influence of reaction time on subjective cognitive load. Regarding MF and SEX, the overall TMS score is analyzed as a measure for state mindfulness.

5.3.9 Data Processing

For the behavioral data, outliers were determined by a deviance of more than three standard deviations from the mean reaction time of all stimulus pairs with the same rotation angle and were excluded from all analyses. Because angular disparity is not defined for mirrored responses in cube figures (Jolicœur et al., 1985; Shepard & Metzler, 1971), only non-mirrored stimulus pairs were analyzed and reaction time was additionally only analyzed for correct responses. The data was further processed in R (version 3.5.1; R Core Team, 2018) to obtain valid data. As it was an online experiment with a twenty-minute intervention, the data had to be scrutinized, processed, and filtered regarding participants' compliance and other factors to accomplish valid data. As such, the data were examined and filtered according to predefined guidelines regarding time-stamps and noticeable behavioral data anomalies.

As the last question of the questionnaire, the participants were asked to declare honestly whether they had been attentive and participated committedly throughout the experiment. Accompanying this question, the information was given that the credits for participation would be attained in any scenario. As a result, 36 participants declared that they had occupied themselves with other things such as using their mobile phones or cooking, or that they had fallen asleep during the experiment. To examine the remaining sample, all actions throughout the experiment were time-stamped. Using this information, an additional filter was a time window of 30 minutes between the end of the pretest and the beginning of the posttest. Within this time frame, the 20-minute intervention was conducted, and thus, whoever took longer to continue the experiment by commencing with the posttest was filtered. As a result, six participants were filtered. Additional filters and respective inspection of the data led to two more participants to be filtered. One took longer than 80 minutes from the beginning of the pretest until the end of the posttest, and showed response times in the mental rotation tasks constantly over 30 seconds. The other participant had clicked through the tests as fast as possible with reaction times of 10 – 300 ms, resulting in a total time of 28 minutes. Further filters regarding the behavioral data led to 14 more participants to be excluded from statistical analysis. For instance, one exclusion criterion was having mean accuracy values of less than 80% for non-rotated and non-mirrored figures. Moreover, nine participants had accuracy means around 50% and overall short reaction times. Another example is one person, who had less than 300 ms and bad accuracy for all figures rotated at 180°, which might indicate frustration and the tendency to just skip through the experiment.

5.3.10 Statistical Analysis

Statistical analysis was performed according to the first study of this thesis (Bauer et al., 2022), and Jost et al. (2023) using *Linear Mixed-Effects Models* using the *lme4* package (version 1.1-21; Bates, Maechler, et al., 2015). Reaction time and cognitive load were analyzed using linear mixed models, and accuracy was analyzed using generalized linear mixed models with a binomial distribution. Model parameters were estimated by maximum likelihood estimation using the *bobyqa* algorithm wrapped by *optimx* package (version 2018-7.10; Nash & Varadhan, 2011) as optimizer. Model fit was calculated using likelihood ratio tests to compare models with and without the fixed effect of interest. The resulting *p* values were compared to a significance level of .05. For multiple comparisons of the same variables, the significance level was Bonferroni corrected. Visual inspection of residual plots did not reveal deviations from homoscedasticity or normality in any model. For the significant effects and main effects, we report the unstandardized effect sizes, and the confidence intervals that were calculated using parametric bootstrapping with 1000 simulations, in line with recommendations of Baguley (2009), and Pek and Flora (2018).

Subjective cognitive load analysis was additionally conducted using *Cumulative Link (Mixed) Models* using the *Regression Models for Ordinal Data* package (*ordinal*, version 2019.12-10; Christensen, 2015). This was done to confirm the results analyzing the cognitive load values rated on a Likert-type scale on an ordinal level. This is based on an ongoing debate whether ratings of perceived exertion scales or similar Likert-type scales should be analyzed as either interval or ordinal data (see Bishop & Herron, 2015). Since both analyses of cognitive load resulted in the same model, we report the results of the *Linear Mixed-Effects Model* for better comparability with the other results.

Model building was based on the research of Barr et al. (2013), and Bates, Kliegl, et al. (2015), starting with a model with random intercepts and slopes for every appropriate fixed effect and reducing the model complexity by dropping non-significant variance components. Non-significant fixed effects were further removed stepwise from the model, that is, effects that least decreased the model fit were removed first, and a model only containing significant fixed effects remained. Then, non-significant effects were tested for improving the model fit by including them in the resulting model. Also, significant effects were tested for worsening the model fit by exclusion of the effect. Main effects for significant interactions were tested separately by splitting the interaction (see also Jost & Jansen, 2020). The resulting models for each parameter are described in the results section. In the tables, all results (i.e., partial interactions, test statistics, and confidence intervals) are depicted for effects with $p < .05$. For $.1 > p > .05$, the test statistics are also depicted. Furthermore, for $p > .1$, only the test statistics for hypothesis-relevant effects (interactions of MF, PP, SEX, and DEG, and selected main effects) are depicted. That means, all other effects in the models resulted in $p > .1$. The TMS scores were analyzed with Mann-Whitney tests in R (R Core Team, 2018). All data were visualized using *ggplot2* package (Wickham, 2016) in R (R Core Team, 2018).

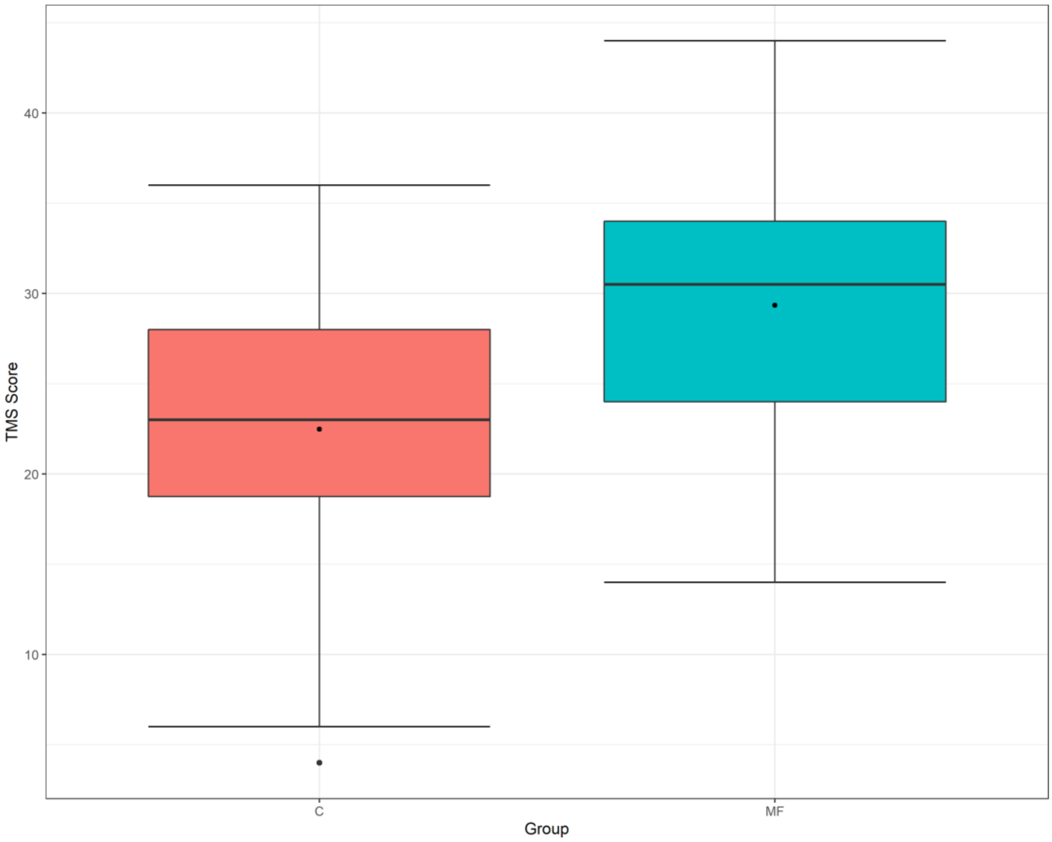
5.4 Results

5.4.1 Toronto Mindfulness Scale Scores

As a measure of state mindfulness, the TMS scores were analyzed. As depicted in Figure 9, the control groups differ visually from the MF groups. This was confirmed by a Mann-Whitney test, which indicated that participants' scores without MF (Mdn = 23) differed significantly to those with MF (Mdn = 30.5), U = 499, $p < .001$. A similar test revealed no significant difference between males (Mdn = 24.5) and females (Mdn = 28), U = 1102, $p = .267$.

Figure 9

TMS Score Plotted by Condition.



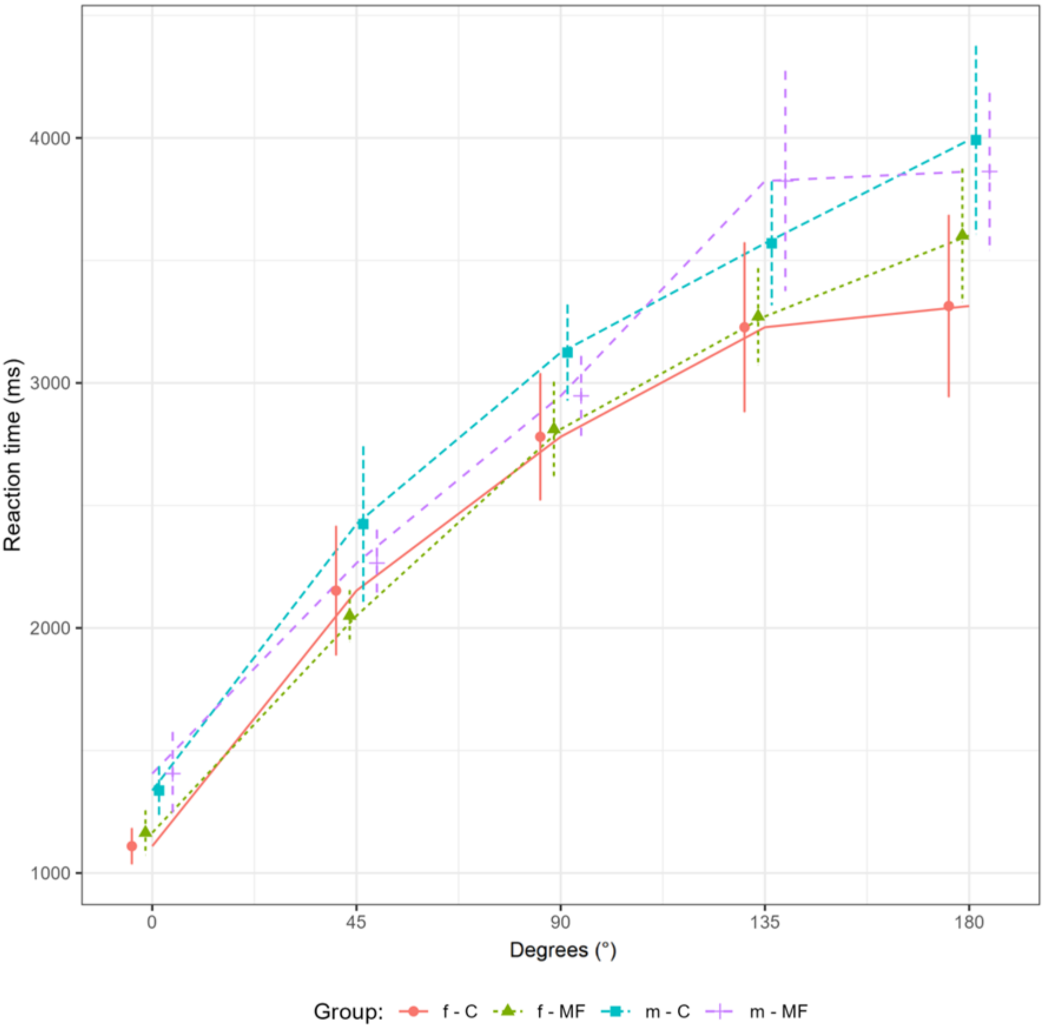
Note. C = control condition, MF = mindfulness condition. The boxes depict the Interquartile Range of the 75th percentile (upper) and 25th percentile (lower) of the data, the whiskers depict the maximum (upper) and minimum (lower) score (excluding outlier), the black lines depict the median, the dots depict the mean (and outlier).

5.4.2 Reaction Time

As shown in Figure 10, the graphs of the reaction times are closely related. All graphs show a positive slope with increasing angular disparity. Visually, the graphs of females are consistently lower than the ones for males, being more separated at higher angular disparity.

Figure 10

Reaction Time of the Posttest Plotted Against Angular Disparity and Group.



Note. f = females, m = males, C = control condition, MF = mindfulness condition. Dots represent the mean. Bars represent the standard deviation. Both are slightly offset for the groups for better visibility.

For reaction time, model construction resulted in a model with random intercepts and slopes for PP and DEG by participant, and for DEG by model. PP*MF*SEX*DEG and all respective interactions and main effects were analyzed as fixed effects. Significant differences were found for PP*MF*SEX and DEG*SEX (see Table 6). The interaction DEG*SEX shows a significant increase in reaction time with increasing DEG, with higher increases for males than females. Reaction time increased significantly by DEG, and decreased significantly by PP (main effects). Figure 11 illustrates the interaction of PP*MF*SEX. All groups improved from pre- to posttest. Males had overall higher average reaction times than females. Visually, for males, the mindfulness group improved less than the control group, and vice versa, for females, the mindfulness group improved more than the control group. Breaking this interaction into its subparts resulted in no significant interactions or main effects.

Table 6

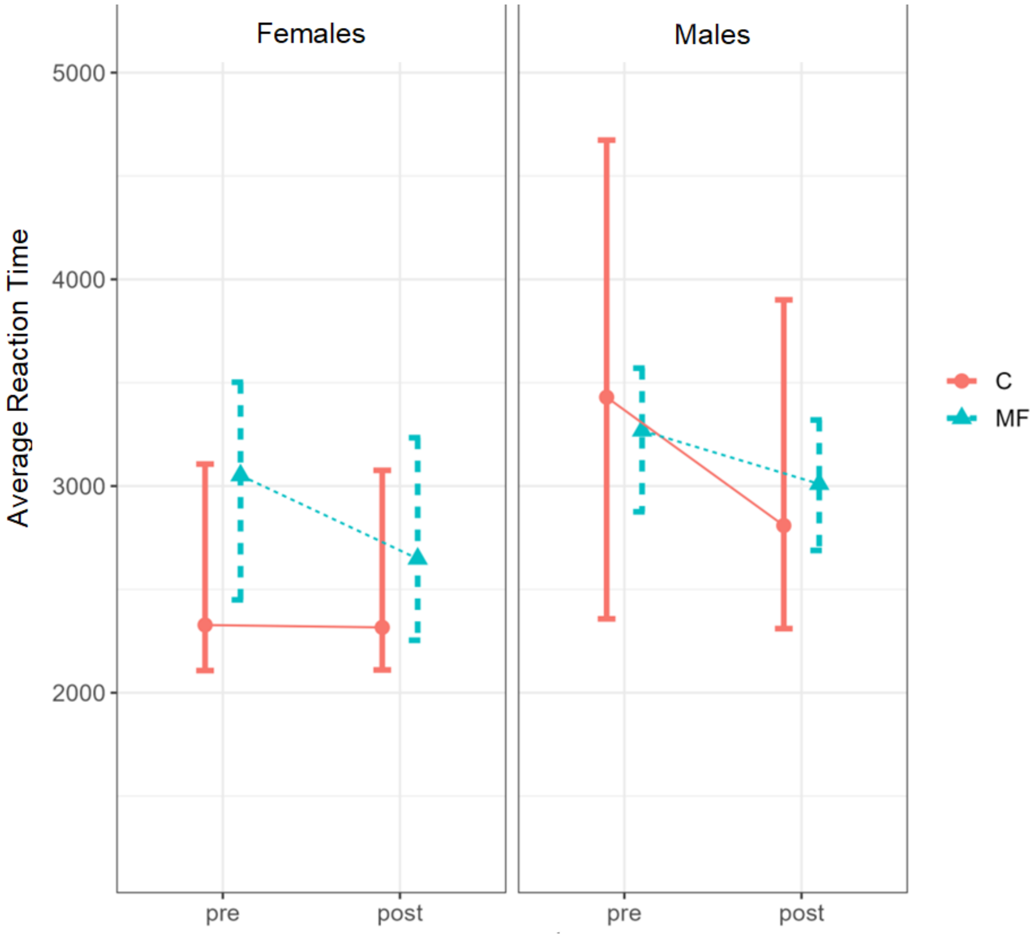
Statistical Analysis of Reaction Time (in Seconds).

Variable	Estimate	SE	Test Statistic	*p* value	95% CI
Intercept	1.49	0.13			1.25, 1.73
MF*PP*SEX*DEG			$\chi^2(1)=0.01$	.974	
MF*PP*DEG			$\chi^2(1)=1.18$	.279	
MF*PP*SEX	0.43	0.20	$\chi^2(1)=4.58$	.032	0.02, 0.81
MF*PP			$\chi^2(1)=0.17$	.677	
PP*SEX			$\chi^2(1)=0.09$	.869	
MF*SEX			$\chi^2(1)=0.11$	.735	
DEG*PP			$\chi^2(1)=0.13$	.719	
DEG*SEX			$\chi^2(1)=4.16$	.041	
DEG*SEX(m)	1.40	0.11			1.18, 1.62
DEG*SEX(f-m)	-0.26	0.13			-0.53, -0.08
DEG(0°)*SEX(f-m)	-0.03	0.17			-0.39, 0.32
DEG(100°)	1.28	0.08	$\chi^2(1)=35.17$	<.001	1.09, 1.46
PP(pretest)	0.28	0.09	$\chi^2(1)=7.57$	.006	0.09, 0.46

Note. Intercept in this model represents the estimate for SEX = males, angular disparity (DEG) = 0°, pre-posttest (PP) = post, and no mindfulness (no MF). Effect of SEX represents the difference between females (f) and males (m). Effects of angular disparity (DEG) represent changes of 100°.

Figure 11

Average Reaction Time Plotted For Sex, Condition, and Pre-Post Test.



Note. C = control condition, MF = mindfulness condition, pre = pretest, post = posttest. Whiskers depict the Interquartile Range of the 75th percentile (upper) and 25th percentile (lower) of the data; the dots connected with a ghost line depict the median.

The results did not support our first hypothesis that mental rotation task performance would be better for the mindfulness condition in the posttest (PP*MF *ns*), and the effect of this was neither modified by increasing angular disparity (PP*MF*DEG *ns*).

Regarding reaction time and our other hypotheses, sex differences did not emerge for the main effect (SEX), but for the interaction PP*MF*SEX, indicating different improvements through the mindfulness induction in interaction with sex and pre-post testing. The results were in line with the mental rotation paradigm, showing increasing reaction times with increasing angular disparity. Our hypothesis that higher improvements would show for larger angles in the posttest was not confirmed (PP*DEG *ns*). However, PP and DEG both showed significant main effects, indicating an overall improvement from pre- to posttest, as well as the general negative effect of increasing angular disparity on reaction times. Additionally, the interaction DEG*SEX indicates that over all tests, increases in angular disparity affected the performance of males significantly more than females.

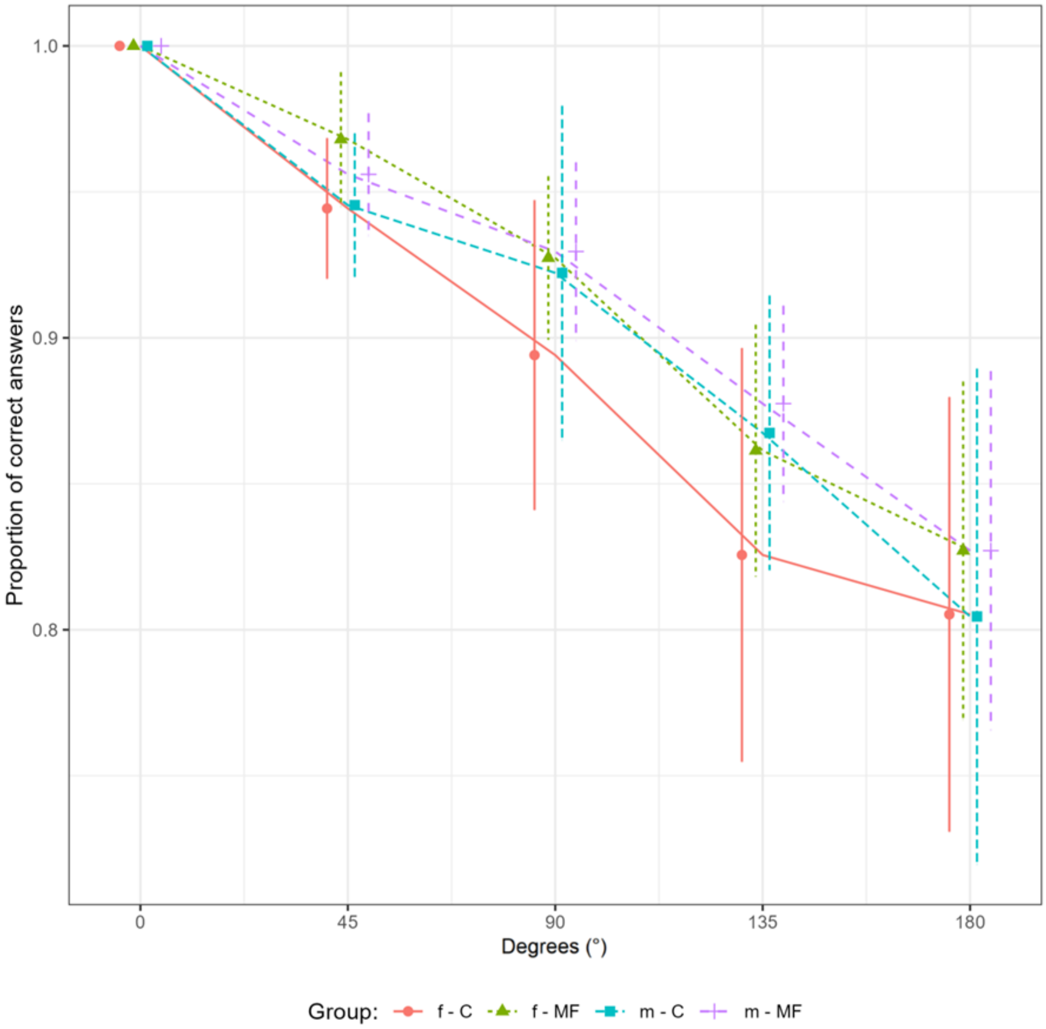
5.4.3 Accuracy

In Figure 12, the graphs of the proportion of correct answers are closely related. All graphs start similarly at zero degrees and show a negative slope with increasing angular disparity. Visually, the graph of the female control group is lower than the other graphs for 90 and 135 degrees of angular disparity. Also, visually the graphs of the control groups tend to be slightly lower than the mindfulness groups.

Similar to reaction time, model construction for ACC resulted in a model with random intercepts and slopes for PP and DEG by participant, and for DEG by model. PP*MF*SEX*DEG and all respective interactions and main effects were analyzed as fixed effects. Significant differences were only found for DEG and PP (see Table 7). Accuracy decreased significantly by DEG, and increased significantly by PP.

Figure 12

Accuracy of the Posttest Plotted Against Angular Disparity and Group.



Note. f = females, m = males, C = control condition, MF = mindfulness condition. Dots represent the mean. Bars represent the standard deviation. Both are slightly offset for the groups for better visibility.

Table 7

Statistical Analysis of (Logarithmic Odds of) Accuracy.

Variable	Estimate	SE	Test Statistic	*p* value	95% CI
Intercept	3.89	0.17			3.54, 4.26
MF*PP*SEX*DEG			$\chi^2(1)=0.12$	.725	
MF*PP*DEG			$\chi^2(1)=0.08$	.781	
MF*PP*SEX			$\chi^2(1)=0.47$	.496	
MF*PP			$\chi^2(1)=0.01$	.969	
PP*SEX			$\chi^2(1)=0.04$	.848	
DEG*SEX			$\chi^2(1)=1.04$	.308	
DEG(100°)	-1.26	0.10	$\chi^2(1)=31.43$	<.001	-1.45, -1.06
PP(pretest)	-0.45	0.18	$\chi^2(1)=4.86$	.028	-0.83, -0.07

Note. Intercept in this model represents the estimate for the logarithmic odds for SEX = males, angular disparity (DEG) = 0°, pre-posttest (PP) = post, and no mindfulness (no MF). Effect of SEX represents the difference between females (f) and males (m). Effects of angular disparity (DEG) represent changes of 100°.

Similar to reaction times, the results for accuracy did not support our first hypothesis that mental rotation task performance would be better for the mindfulness condition in the posttest (PP*MF *ns*), and the effect of this was neither modified by increasing angular disparity (PP*MF*DEG *ns*).

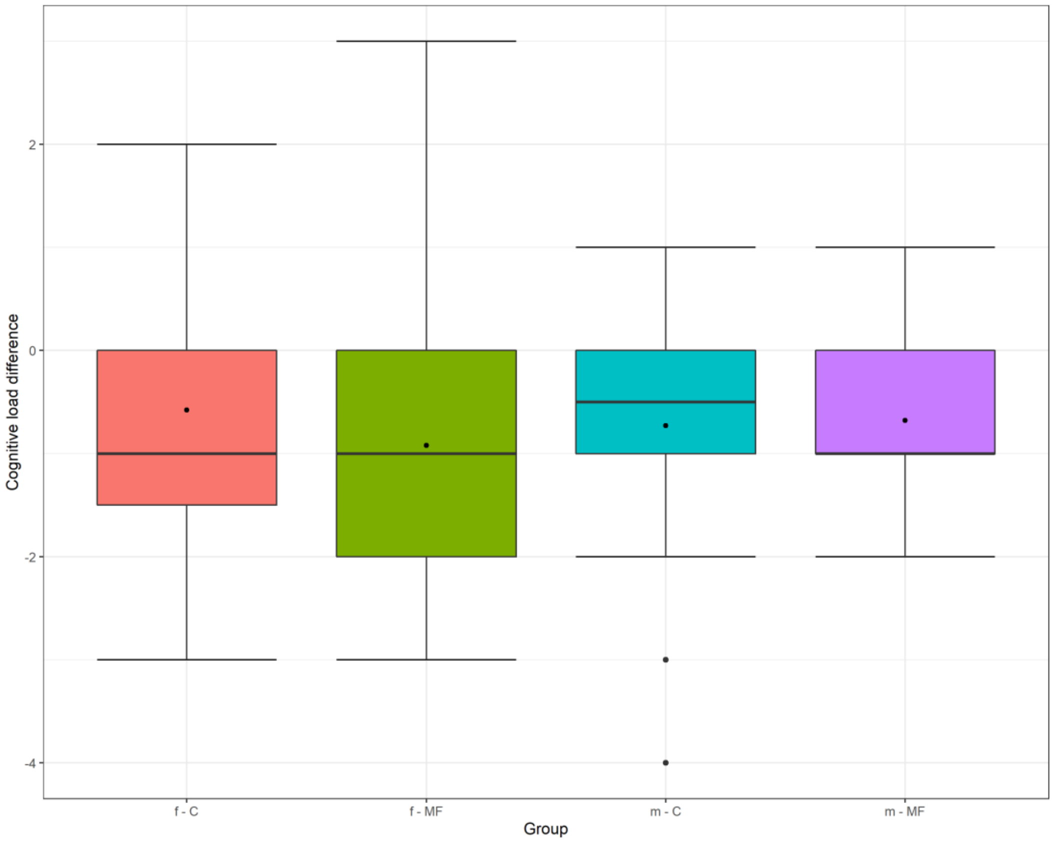
Regarding accuracy and our other hypotheses, sex differences emerged neither for the main effect (SEX) nor for the interactions. The results were in line with the mental rotation paradigm, showing decreasing accuracy rates with increasing angular disparity. Our hypothesis that higher improvements would show for larger angles in the posttest was not confirmed (PP*DEG *ns*). However, PP and DEG both showed significant main effects, indicating an overall improvement from pre- to posttest, as well as the general negative effect of increasing angular disparity on accuracy rates.

5.4.4 Subjective Cognitive Load

The differences of the (self-reported) subjective cognitive load are depicted in Figure 13 and indicate an overall reduction of perceived cognitive load in the posttest for all groups. The groups show similar values, with smaller dispersion for males.

Figure 13

Changes of Subjective Cognitive Load (Posttest – Pretest Values) Plotted by Group.



Note. f = females, m = males, C = control condition, MF = mindfulness condition. The boxes depict the Interquartile Range of the 75th percentile (upper) and 25th percentile (lower) of the data, the whiskers depict the maximum (upper) and minimum (lower) score (excluding outlier), the black lines depict the median, the dots depict the mean (and outlier).

The model building for cognitive load resulted in a model with random intercepts by participant. PP*MF*SEX, all respective interactions and main effects, and RT-Sum were analyzed as fixed effects. The cognitive load decreased significantly by PP, and increased significantly by RT-Sum (see Table 8). With regard to the inclusion of RT-Sum in the model, we inspected the data for possible collinearity problems. All variance inflation factors were smaller than three (maximum of 1.03), that means, collinearity was not an issue (see Zuur et al., 2010).

Table 8

Statistical Analysis of Subjective Cognitive Load.

Variable	Estimate	SE	Test Statistic	*p* value	95% CI
Intercept	3.69	0.26			3.19, 4.17
MF*PP*SEX			$\chi^2(1)=0.16$	.691	
MF*PP			$\chi^2(1)=0.54$	.463	
PP*SEX			$\chi^2(1)=0.28$	.596	
PP(pretest)	0.67	0.13	$\chi^2(1)=24.37$	<.001	0.44, 0.94
RT-Sum	0.08	0.03	$\chi^2(1)=6.75$	.009	0.02, 0.14

Note. Intercept in this model represents the estimate for SEX = males, pre-posttest (PP) = post, and no mindfulness (no MF). Effect of SEX represents the difference between females (f) and males (m).

With the variable RT-Sum having an impact on cognitive load, we exploratively analyzed this variable more. It is important to note that this analysis is different from the analysis of reaction times by item, which is only conducted for non-mirrored and correctly answered ones. The total sum of reaction time describes the amount of time per block (pre- or posttest), in which participants needed to respond to all 150 items. Statistical analysis showed significant main effects for PP and SEX. Both sexes have smaller sums of reaction times in the posttest than pretest. Additionally, males have overall higher values than females in both pre- and posttest. For this analysis, collinearity was not a problem, as all variance inflation factors were smaller than one (see Zuur et al., 2010).

In the results for subjective cognitive load, there was no evidence for our second hypothesis that self-reported cognitive load would be influenced by the mindfulness induction and be lower in the posttest (PP*MF and MF *ns*). Our third hypothesis was confirmed, as subjective cognitive load decreased overall from pre- to posttest (PP sign.). In our fourth hypothesis, we expected subjective cognitive load for pre- and posttests to be linked to the overall sum of reaction times. This was confirmed, as increasing overall response time led to higher values in subjective cognitive load (RT-Sum sign.).

Regarding subjective cognitive load and our other hypotheses, no sex differences emerged neither for the main effect (SEX) nor for the interactions. In the explorative analysis of the variable RT-Sum, the sums of reaction times decreased from pre- to posttest for all groups, and a significant main effect of sex indicated that females had overall lower reaction time sums in both tests.

5.5 Discussion

5.5.1 Toronto Mindfulness Scale Scores

All participants had to fill out a TMS questionnaire. The TMS scores were then analyzed to determine whether the mindfulness induction had the wanted effect. The participants of the mindfulness groups showed overly high scores for state mindfulness in the experiment (see Lau et al., 2006). These scores are similar to the ones Weger et al. (2012) reported for their mindfulness groups. The participants of the control groups showed significantly lower scores. Therefore, the twenty-minute mindfulness induction elicited a significant main effect on state mindfulness.

5.5.2 Mental Rotation Performance

Based on our hypotheses, the focus was on the interaction between mindfulness and pre-post testing, which would illustrate the changes in behavioral performance resulting from the treatment. As we expected these changes to be larger for tasks of higher difficulty, we added angular disparity to the modeling. In addition, we examined exploratively whether sex had an influence.

In the analysis for reaction time, the interaction between pre-post testing, mindfulness, and sex was significant. This indicates that in interaction with the mindfulness induction or control condition, men and women showed significantly different changes between the pre- and posttest. Visual inspection of the reaction time averages (see Figure 11) might indicate that men improved more in the control condition than in the mindfulness one, and that women improved more in the mindfulness condition than the control one. Dividing this three-way interaction into its subparts resulted in no significant effects. Due to the filtering of many participants, the explorative analysis of this three-way interaction was underpowered. The effect size of this interaction could be taken into account for future studies for sample size estimations.

However, another interaction was significant, that is, sex interacted with changes of angular disparity. This means that in both tests, reaction time increased more for men with higher levels of angular disparity than it did for women. Hence, in both conditions, women seemed to be less susceptible to the higher difficulty due to larger rotation angles between the figures.

For accuracy, no interactions between pre-post testing and mindfulness emerged as significant, indicating that no effects regarding the mindfulness induction can be observed. In the model, angular disparity and pre-post testing were the only significant variables. This could be explained by the actuality that in mental rotation tasks, a speed-accuracy-tradeoff always happens.

Here, accuracy undergoes smaller changes than reaction time and is therefore less influenced by different variables (see e.g., Hertzog et al., 1993; Wickelgren, 1977).

With regard to our first hypothesis, no significant differences in the performance manifested between the two treatment conditions. Only with the addition of sex in the interaction, the groups differed significantly in the model. This is interesting, as other chronometric mental rotation studies (e.g., Jansen-Osmann & Heil, 2007; Voyer et al., 2006) reported no sex differences in mental rotation performance of 3D cube figures. Considering the effective change in state mindfulness, the embodied meditation treatment could have had more impact on attentional control for the women, whereas the men may have profited more from the control condition (see Lutz et al., 2008). This could mean that the embodiment effects were less beneficial for the men. However, as no further interactions or main effects regarding these interactions with sex emerged, the conclusions of this three-way interaction are limited. Future research could analyze this further to determine more details.

Altogether, our results did not confirm our hypothesis that task performance would be better after the mindfulness induction. Within the three-way interaction, there is only a visual indication that women might have benefited more from the mindfulness induction than men did. This stands in contrast to the findings of Weger et al. (2012), on which our hypothesis was based. In the study with mathematical tests, the authors reported a significant main effect of a brief five-minute mindfulness intervention (eating two raisins) versus a control condition. Similarly, in their recent study with psychometric mental rotation tasks, Rahe and Jansen (2023) reported a significant effect of a five-minute mindfulness intervention (eating two candies). As our intervention was recorded by a professional instructor and lasted longer (i.e., twenty minutes) than the ones Weger et al. (2012), and Rahe and Jansen (2023) induced (i.e., five minutes), and with the TMS scores demonstrating a significant difference in elicited state mindfulness in the intervention group compared to the control group, the different results regarding the (non-

)existence of mindfulness effects are noteworthy. These differences might not only derive from the different treatment durations in the studies, but also the different forms of focused attention meditation. The body-scan and the raisin-/candy-eating both involve aspects of embodiment. According to the embodied cognition approach, they would thus elicit respective embodiment effects. By design, the body-scan incorporates more the whole body and the sensations regarding several body areas. In contrast, the eating meditations involve more focus on the hands and the mouth, and respectively touch and taste. This tactile and taste sensory information could lead to more pronounced or different embodiment effects, which seem to influence cognitive performance more impactfully. As brief mindfulness inductions can influence the working memory (e.g., Beilock et al., 2007), the effects may be more pronounced for the eating than the body-scan treatment.

In line with the paradigm for chronometric mental rotation tasks, changes in angular disparity significantly influenced behavioral data. Higher angular disparity between the two pictures resulted in higher reaction times and lower accuracy. Differences between the reaction times of pre- and posttest can be observed, with worse performance values in the pretest. These outcomes can be explained as typical learning effects in a pre-post experimental design.

5.5.3 Subjective Cognitive Load in the Mental Rotation Task

Regarding the analysis of subjective cognitive load, in our third hypothesis, we expected reduced load for the mindfulness condition in the posttest. This hypothesis was based on the effects of embodied mindfulness (such as the body-scan meditation used here) to influence the brain's introspective systems (underlying conscious experience of emotion, motivation, and cognitive operations) to elicit improvements in cognitive tasks (see Niedenthal et al., 2005) and reduce subjective cognitive load. However, our analysis showed no evidence in support of this hypothesis. The mindfulness induction and the control condition did therefore not differ significantly in their effect on subjective cognitive load. However, for all groups, the subjective cognitive load decreased from the pretest to the posttest, confirming our third hypothesis. First, this result could be explained by learning and habituation effects, which could lead participants to ascertain task difficulty and resulting cognitive strain to be lower in the second mental rotation part of the experiment. Second, both conditions involved listening to a voice for twenty minutes. While the TMS scores indicate that there was a difference in elicited state mindfulness induction between the condition groups, possible effects of just concentrating and listening to a recording can still have an influence for both groups. For instance, participants could have experienced both conditions to calm the mind and thus reduce arousal. This notion could be supported by the findings of Jürgens et al. (2018). In their study, they measured arousal differences via pupillometry for different levels of affective prosodic utterances. In the presentation of neutral prosodic stimuli, the pupil sizes were significantly lower than for joyful or angry ones. Similarly, in our study, the control condition was chosen to be neutral, which could mean that the control condition could lead to lower levels of cognitive effort. Thus, aside from the other effects of the embodied meditation practice, both conditions could therefore have led to lower overall estimations of subjectively perceived cognitive load, whose range of possible differences resulted as non-significant.

In our fourth hypothesis, we expected participants' subjective cognitive load for pre- and posttests to be linked to their overall sums of reaction times. The significant main effect of the reaction time sums confirmed this hypothesis. This means that with increases in the summed reaction times, participants would rate their subjective cognitive load higher. As this variable was only measured twice, that is, every time after completing the whole test block in pre- and posttest, the longer perceived time to complete these blocks seems to influence the subjective rating of cognitive load. This could also be in part due to familiarization with the mental rotation test. Notably, changes of subjective cognitive load did not get fully explained by effects of pre-post testing, but the sums of the reaction times still predicted a significant portion of the changes of subjective cognitive load in the statistic model.

As the sums of all the reaction times resulted to have a significant effect, we further analyzed this variable exploratively. In the analysis, these sums decreased from pre- to posttest for all groups, also illustrating possible learning effects. Furthermore, a significant main effect of sex emerged, indicating that women had overall lower sums of reaction times in both tests. This is interesting, as no main effect of sex differences emerged in any other analysis of the dependent variables. One minor difference is only the lower variance in the data of cognitive load for men.

In the analysis of sex differences for subjective cognitive load, no interaction or main effects were significant. Overall, both sexes rated the pre- and posttest similarly. Men only show a smaller variance compared to women. As part of their study on the link between mental rotation and physical exercise, Jost et al. (2023) analyzed subjective cognitive effort for mental rotation task blocks. In their analysis with 22 women and 19 men, they found a significant main effect of sex, with subjective cognitive effort being significantly lower for women than men. On the other hand, in the pupillometric study of Campbell et al. (2018), women showed higher cognitive load—or used more attentional resources—than men in mental rotation tasks of abstract figures. Taking into consideration that the measures of cognitive load differ between this study and Campbell et al. (2018), our findings did neither confirm their results, nor the findings of Jost et al. (2023), as no differences emerged. As a conclusion, both sexes solved mental rotations tasks with cube figures of similar difficulty and rated their subjective cognitive effort similarly. Our last study showed corresponding results regarding the sex differences in cognitive effort objectively measured via pupillometry (see the first study of this thesis; Bauer et al., 2022). Hence, our findings in this study provide further evidence that men and women perceive (and probably need) the same cognitive load for the solution of cube figure tasks. These possible effects will be analyzed further in **Study 3** of this thesis, including an interaction with stereotype threat application and analysis of objective measures of cognitive load via pupillometry.

5.6 Limitations

Instead of the option to measure subjective cognitive load after every item, we chose to only measure it after each block. On the one hand, the results are less detailed and imply a greater internal variance for each participant. On the other hand, the common chronometric mental rotation task design was maintained and the overall experiment duration was kept to a reasonable participation time. However, for future research, more detailed measurements per item could be considered.

One more limitation is the short mindfulness induction. As it was the objective to research the effects of a single bout of brief mindfulness induction, it is still important to consider that mindfulness general works best through daily practice routine over a longer period of time (Rahe & Jansen, 2023).

As mentioned in the discussion, the filtering in the preprocessing led to a final sample size that was smaller than estimated beforehand, resulting in an underpowered explorative analysis of sex differences.

5.7 Conclusion

In this online study, the twenty minute mindfulness induction resulted in elevated state mindfulness for the mindfulness condition, which was significantly different to the control condition. The main effects of the mindfulness induction were not significant in any of the dependent variables; however, an interesting three-way interaction emerged between pre-post testing, mindfulness, and sex for reaction times. This interaction indicated that women might have benefited more from the mindfulness induction and men more from the control condition. Additionally, the models showed that men's performance was more affected by increasing levels of angular disparity than women's in pre- and posttest. Analysis of subjective measures of cognitive load showed reduced load in the posttest as well as a connection to reaction times. In the model, larger sums of reaction times would lead to significantly higher levels of subjective cognitive load. This variable was analyzed exploratively and showed sex differences, with men taking significantly longer than women. As this was not the main interest of this study, this variable could be analyzed further in future research.

6 Third Study: The Effect of Mindfulness and Stereotype Threat in Mental Rotation: A Pupillometry Study[2]

6.1 Summary

In this study, the effects of a brief mindfulness intervention and a stereotype threat activation on behavioral performance and cognitive load in a chronometric mental rotation test with cube figures were investigated. 107 participants (55 women and 52 men) were divided into four test groups (mindful or control intervention combined with or without stereotype threat activation). They completed two sets of 150 items each with cube figures of similar complexity and filled out a test of state mindfulness. Cognitive load was measured by changes in pupil dilation using eye tracking. Cognitive load, reaction time, and accuracy were analyzed with linear mixed models with the factors time, stereotype threat, mindfulness, angular disparity, and sex. The analyses revealed no sex differences in either of the measurements. Neither stereotype threat nor mindfulness influenced task performance significantly. However, we found that high levels of state mindfulness inhibit stereotype threat to negatively influence task performance and cognitive load.

6.2 Goals and Hypotheses

With mathematical abilities and mental rotation performance being related (Xie et al., 2020) and mental rotation also being partially stereotyped (e.g., Heil et al., 2012; Moè & Pazzaglia, 2006), the reported effects of Weger et al. (2012) in mathematics might also apply to mental rotation tasks. Our goal was to conduct a study designed in close relation to Weger et al. (2012). In their study, they tested the effects of stereotype threat on women (N = 71) in a 2x2 mixed design. Similarly, our participants performed mental rotation tasks and were also assigned to four

---

[2] The results presented in this chapter were published in advance in: Bauer, R., Jost, L., & Jansen, P. (2021). The effect of mindfulness and stereotype threat in mental rotation: a pupillometry study. *Journal of Cognitive Psychology, 33*(8), 861-876.

condition groups with or without mindfulness intervention and with or without stereotype threat. In addition to this, we also tested men to compare the results for both sexes. Even more, we used pupillometry to investigate if the sex differences in mental rotation are due to cognitive load. The following hypotheses were investigated:

In line with the findings of Weger et al. (2012), we predicted an interaction between stereotype threat and mindfulness, with mindfulness only improving the mental rotation performance in the stereotype threat condition for women (Hypothesis 1). We expected these changes to be larger for tasks of higher difficulty, that is, a higher angular disparity between both objects being compared. Mindfulness together with stereotype threat was expected to reduce cognitive load by showing lower pupil dilation in the posttest (Hypothesis 2). Stereotype threat by itself was expected to increase cognitive load (Hypothesis 3). In form of additional hypotheses, we expected an influence of reaction time on pupil dilation due to their close relation indicating cognitive effort. Furthermore, we expected a time effect of angular disparity, that is, greater performance improvements in the posttest for higher angular disparity. The effect for sex differences will be analyzed, too, exploratively.

6.3 Methods

6.3.1 Participants

In the study, 125 students (61 women) participated and received study credits. Eye tracking recording errors led the data of 18 participants to be excluded from further analysis. Consequently, 107 students (55 women, mean age (SD) = 20.8 (1.8) years; 52 men, mean age (SD) = 21.4 (2.2) years) form the sample for statistical analysis. In the study of Weger et al. (2012), the interaction effect of mindfulness on stereotype threat for women was detected with a $d = .92$. The required sample size for women was estimated using the software G*power (Faul et al., 2007). To ensure a sufficient sample size, the effect size of $d = .92$ ($f = .46$), a power of 0.90, and an alpha value of .05, a sample size of 52 women was estimated for this study (see Brysbaert & Stevens, 2018). For participation, unrestricted eyesight at close range or corrected eyesight with contact lenses was required, as was no practical experience with mindfulness. Thirty participants reported minor theoretical knowledge with mindfulness, and the others reported no knowledge of mindfulness at all. Participants were sports and movement sciences students, of which 68 reported no experience with mental rotation tasks, 19 reported having participated in one mental rotation task experiment, and 20 reported having participated in two or three mental rotation task experiments prior to this experiment. In total, the test sample had an average of 0.61 mental rotation task experiment experience prior to this study. To mitigate possible experience effects, participants were required not to have participated in any mental rotation task experiment within the six months preceding this experiment. Participants were free from eye injuries and reported no relevant physical or mental limitations. Informed consent was obtained from all individual participants. The experiment was conducted according to the ethical declaration of Helsinki. We communicated all considerations necessary to assess the question of ethical legitimacy of the study.

6.3.2 Conditions

Participants were assigned to one of four test groups in a counterbalanced order. The intervention after completing the first block of mental rotation tasks depended on this assigned group. This intervention was either a five-minute mindfulness intervention, during which the participants listened to recorded instructions by a professional *Mindfulness-Based Stress Reduction* instructor to eat two raisins in a mindful manner, or a control intervention, during which the participants were instructed to eat two raisins in a five-minute break sitting silently in their chair. Following this, participants either did or did not receive a stereotype threat condition from the investigator. The stereotype threat consisted of informing the participant that he or she was taking part in an experiment about "why women/men are always better in solving mental rotation tasks" (always using the respective counterpart to induce the stereotype threat). This was similar to the approach of Weger et al. (2012) to present a standardized brief stereotype threat in one sentence (see also, Colzato et al., 2016). Thirteen men were assigned to each of the four conditions (control/noST, control/ST, mindfulness/noST, or mindfulness/ST). Thirteen women were assigned to the control condition without stereotype threat, and 14 to each of the other three.

6.3.3 Setup

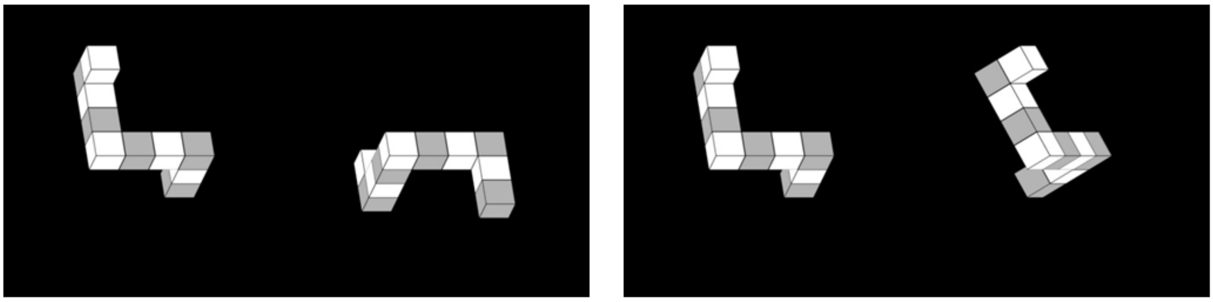
Stimulus presentation and response handling were controlled with Presentation® software (Version 20.1 Build 12.04.17, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com) on a Dell Latitude E5540 Laptop, 15", 1366 x 768 px, 60 Hz. Below the bottom screen border, a RED250mobile (SensoMotoric Instruments GmbH, 2017) eye tracker running at 250 Hz was applied. iViewRED software (Version 4.4.26.0, SensoMotoric Instruments GmbH, 2017) was used to integrate all screen properties and the position of the tracker relative to the screen. With the iViewX_SDK (Version 4.4.10.0, SensoMotoric Instruments GmbH, 2017), the Presentation® script conducted a 13-point calibration at the start of each run. Calibration accuracy in form of vertical and horizontal dispersion was reported to be lower than 0.3° for all participants. The iViewRED software depicts the distance of the eyes to the screen. All participants were seated close to a table while placing the head on a sturdy chinrest. Then, the laptop was positioned with the screen being at 60 cm from participants' eyes. Two tables at different heights were used. On the lower table, participants had their arms and chinrest. The laptop with the eye tracker was placed on the higher table to prevent participants' movement in form of possible small vibrations to corrupt data collection. The participant used a wired mouse for task response to control for input lag. The laboratory was silent and dimly lit to control for pupil dilation due to light conditions. The only light sources were a ceiling lamp outside the peripheral visual field of the participant and the laptop screen (luminance 1.04 cd/m^2 for all model figures; measured with a spot meter, Chroma Meter CS-100, Minolta Co., Ltd., Japan), resulting in a constant illuminance of 22 lx (measured with a lux meter, testo 540, Testo AG, Germany).

6.3.4 Stimuli

Stimuli consisted of a selection from the stimulus library of Peters and Battista (2008). Ten cube figure models with rotations around x- and z- axis in 45° steps and mirrored/non-mirrored orientation (i.e., figures one to ten with *a* and *b* orientation in the stimulus library) were used with a checkered pattern on black background (see Figure 14).

Figure 14

Examples of Mental Rotation Stimuli Used in the Experiment.



Note. Selected from the stimulus library of Peters and Battista (2008). Left: 90° rotation in x-axis. Right: 45° rotation in z-axis.

In the main experiment, which was divided in two parts, each part consisted of 150 stimuli, resulting in a total of 300 different stimuli. The first 150 stimuli were figures one to five, and the second 150 stimuli were figures six to ten of the stimulus library (Peters & Battista, 2008). The order of the stimuli in each block was first randomized once and then counterbalanced in their order over all participants. On the left side of the screen, every model was presented in orientation a, rotated by 30° in x direction and 15° in z direction, so that the base model for x or z rotation was identical. On the right side of the screen, a rotated and mirrored/non-mirrored stimulus was presented. Stimulus pictures were sized 400px times 400px and presented vertically centered and horizontally positioned 300px to the left or right of the center of the screen until a response was given. In the practice block, three different models were presented in overall 24 practice stimuli (figures 11 to 13 of the stimulus library), that is, four rotations in x-axis and z-axis for each model, combined with using each possible angular disparity once and half of the trials being mirrored. In the practice session, between the stimuli pairs, participants received feedback for 1000 ms (✔- right, ✘- wrong) shown at the center of the screen. In experimental sessions, a fixation cross ("+") was shown there for 500 ms.

6.3.5 Pupil Diameter

Before the practice block, baseline measurements were conducted to later calculate comparable pupil diameter data. All ten models were presented in randomized order and only for the 0° non-mirrored condition. Participants were instructed that they were about to see some pictures, which they only had to look at. The illumination of the screen for all individual figures is presented once and can further be used as a controlled baseline for light condition influences (Holmqvist et al., 2011). To address these two aspects of the baseline data, participants looked at each figure for six seconds. For the time interval from three to five seconds after stimulus onset, the average pupil diameter of each eye was calculated for each figure. This time range was chosen to reduce the effects of the onset of new visual stimuli and to control for general physiological variations. The response in pupil dilation typically occurs within the first few hundred milliseconds after stimulus onset, peaks around one to two seconds after stimulus onset and continues asymptotically until it returns to baseline values (Beatty & Lucero-Wagoner, 2000; Nieuwenhuis et al., 2011). For this reason, longer baseline and stimulus durations are recommended, where participants must maintain fixation (e.g., Hayes & Petrov, 2016). In the main experiment, the maximum pupil diameter for each eye was recorded. Then, the difference between this value and its respective baseline value was calculated for each item (Beatty & Lucero-Wagoner, 2000). Based on the fixation duration of each eye, their weighted mean of dilation difference was calculated. Unscrambled pictures were chosen over scrambled ones due to most of the screen (background) being black with the latter, resulting in nearly completely black pictures.

6.3.6 Toronto Mindfulness Scale

As a manipulation check, whether the mindfulness manipulation had the desired effect, we used the Toronto Mindfulness Scale (TMS; Lau et al., 2006) as a measure of state mindfulness. The TMS was translated to German and included 13 items with five answer alternatives (0 = not at all; 4 = very much). For each participant, we computed summed scores for all items (see Lau et al., 2006). To prevent priming the participants to focus on mindfulness behavior prior to the treatment (either control or mindfulness condition), the TMS was only done once after the second mental rotation task block. Thus, these values served as a check, whether the treatments showed an overall different effect between the groups. In this study, the internal consistency of the questionnaire is acceptable, with the Cronbach's alpha of the TMS score being 0.78.

6.3.7 Procedure

The experiment was a single session and lasted between 40 to 50 minutes, depending on participants' speed to complete all items. Upon arrival, the participant was informed about the study protocol (questionnaire, calibration, a mental rotation task block, a short middle part, another mental rotation task block, questionnaire, end) and the *general intention* of the study (i.e., investigation of mental rotation with eye tracking; no information was given about the conditions or pupillometry), then read and signed the informed consent, and filled out a demographic questionnaire. After positioning the participant with the chinrest at the table and a brief explanation of the test protocol, the calibration and baseline were run. The practice session with feedback followed, which was introduced by a digitally presented instruction. Participants used the right hand for mouse handling and received written instructions to press the left mouse button, if the stimuli could be rotated into congruence (non-mirrored, "same"), and the right mouse button, if the two stimuli were mirrored ("different"), and to answer as quickly and precisely as possible. After completing all practice trials, the calibration and the first part of the main experiment were run. Depending on the assigned test group, an intervention (control, control/ST, mindfulness, or mindfulness/ST) followed. After another calibration, the second part of the experiment was conducted. Following this part, the participant filled out the TMS (Lau et al., 2006). In the end, the participant was debriefed. All verbal instructions were standardized.

6.3.8 Study Design

To analyze cognitive performance, the dependent variables are reaction time (RT) and accuracy (ACC). The difference between the maximum pupil diameter and the respective baseline pupil diameter for each trial (PD) serves as a measure of cognitive load. Independent variables are pre-post experiment setup (PP), mindfulness intervention (MF), stereotype threat (ST), SEX, angular disparity (DEG), and their respective interactions. DEG describes the angular disparity between the two figures shown on the screen. Since the left image was always presented with 0° rotation, the angular disparity depicts the rotation in degrees of the right image. The PD analysis includes reaction time as an additional fixed effect. Regarding MF and SEX, the overall TMS score is analyzed as a measure for state mindfulness.

6.3.9 Data Processing

For the behavioral data, outliers were determined by a deviance of more than three standard deviations from the mean reaction time of all stimulus pairs with the same rotation angle and were excluded from all analyses. Because angular disparity is not defined for mirrored responses in cube figures (Jolicœur et al., 1985; Shepard & Metzler, 1971), only non-mirrored stimulus pairs were analyzed, and reaction time was additionally only analyzed for correct responses. Using the SMI software BeGaze 3.7, build 58 (SensoMotoric Instruments GmbH, 2017), the implemented Raw Data Export with a velocity dependent algorithm (peak velocity threshold = 40°/s, min. fixation duration = 50 ms, peak velocity between 20–80% of saccade length) was used to determine fixation, blink, and saccade events.

The data was further processed in R (version 3.5.1; R Core Team, 2018) to obtain valid pupil diameter data. If a trial started with a fixation, its starting time was set to 0 (keeping the pupil diameter). If the ending time of the eye tracking data was longer than the reaction time received from Presentation® for an item, events with the starting time being after the reaction time ending got erased (including the respective pupil diameter). Overlapping events at the trial end got cut to the reaction time (keeping the pupil diameter). If the eye tracking time ended more than 200 ms before the reaction time, the trial was treated as an outlier. The baseline average pupil diameter is the weighted mean (by time) of all event fixations during a two second interval for each eye. The maximum pupil diameter takes the maximum value of all event fixations that have a pupil diameter and is then calculated as a weighted mean (by fixation duration) of the two eyes. In the baseline, the values did neither differ between men (mean (SD) = 4.35 (0.76)) and women (mean (SD) = 4.49 (0.66), $\chi^2(1) = 1.14$, $p = .29$), nor between eyes (left, mean (SD) = 4.43 (0.70); right, mean (SD) = 4.42 (0.72), $\chi^2(1) = 0.93$, $p = .34$), but did differ between models, $\chi^2(9) = 153.69$, $p < .001$. Based on the same reasoning for behavioral data analysis in mental rotation tasks, our pupil dilation analysis was conducted with a dataset filtered similarly to the reaction time dataset, that is, with the maximum pupil diameter excluding mirrored items and wrongly answered ones. This is also in line with the database selection used by Campbell et al. (2018) for pupillometric measurements with mental rotation tasks.

6.3.10 Statistical Analysis

Statistical analysis was performed according to Bauer et al. (2022), and Jost et al. (2023) using *Linear Mixed-Effects Models* using the *lme4* package (Bates, Maechler, et al., 2015; version 1.1-21). Reaction time and pupil diameter were analyzed using linear mixed models, and accuracy was analyzed using generalized linear mixed models with a binomial distribution. Model parameters were estimated by maximum likelihood estimation using the *bobyqa* algorithm wrapped by *optimx* package (version 2018-7.10; Nash & Varadhan, 2011) as optimizer. Model fit was calculated using likelihood ratio tests to compare models with and without the fixed effect of interest. The resulting *p* values were compared to a significance level of .05. For multiple comparisons of the same variables, the significance level was Bonferroni corrected. Visual inspection of residual plots did not reveal deviations from homoscedasticity or normality in any model. For the significant effects and main effects, we report the unstandardized effect sizes, and the confidence intervals that were calculated using parametric bootstrapping with 1000 simulations, in line with recommendations of Baguley (2009), and Pek and Flora (2018).

Model building was based on the research of Barr et al. (2013), and Bates, Kliegl, et al. (2015), starting with a model with random intercepts and slopes for every appropriate fixed effect and reducing the model complexity by dropping non-significant variance components. Non-significant fixed effects were further removed stepwise from the model, that is, effects that least decreased the model fit were removed first, and a model only containing significant fixed effects remained. Then, non-significant effects were tested for improving the model fit by including them in the resulting model. Also, significant effects were tested for worsening the model fit by exclusion of the effect. Main effects for significant interactions were tested separately by splitting the interaction (see also Jost & Jansen, 2020). The resulting models for each parameter are described in the results section. In the tables, all results (i.e., partial interactions, test statistics, and confidence intervals) are depicted for effects with $p < .05$. For $.1 > p > .05$, the test statistics are also depicted. Furthermore, for $p > .1$, only the test statistics for hypothesis-relevant effects (interactions of MF, ST, PP, SEX, and DEG, and selected main effects) are depicted. That means, all other effects in the models resulted in $p > .1$. The TMS scores were analyzed with Mann-Whitney tests in R (R Core Team, 2018). In all graphs, the depicted data are combined for both sexes for better readability. All data were visualized using *ggplot2* package (Wickham, 2016) in R (R Core Team, 2018).

6.4 Results

6.4.1 Toronto Mindfulness Scale Scores

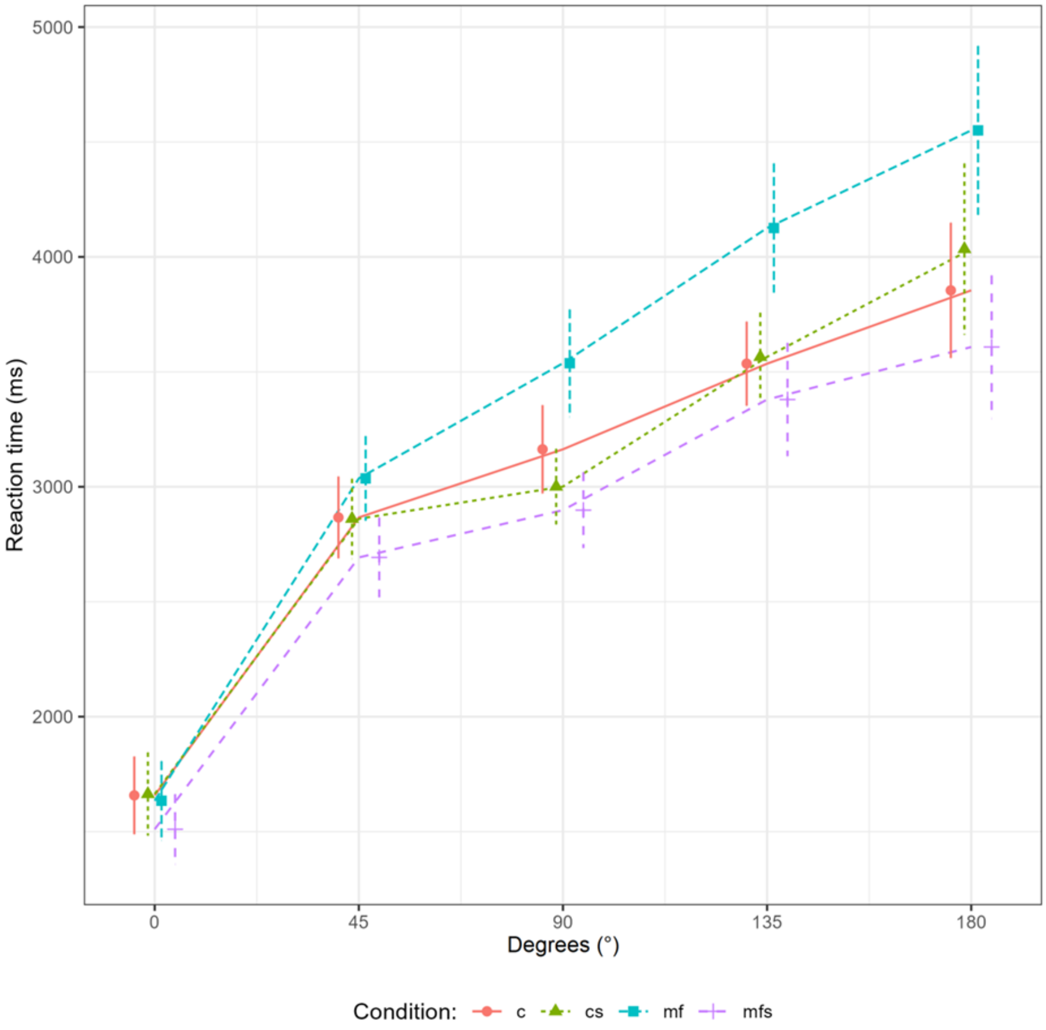
As a measure of state mindfulness, the total TMS scores were analyzed. Visual inspection and test for normality (Shapiro-Wilk) revealed the data to be significantly different from a normal distribution. A Mann-Whitney test indicated that participants' scores with MF (Mdn = 30.5) did not differ significantly to those without MF (Mdn = 27), U = 1693.5, $p = .102$. A similar test also revealed no significant difference between men (Mdn = 28) and women (Mdn = 29), U = 1473, $p = .791$.

6.4.2 Reaction Time

As shown in Figure 15, the graphs of the reaction times are closely related, with the mindfulness group visually showing slightly higher values for higher angular disparity. All graphs show a positive slope with increasing angular disparity.

Figure 15

Reaction Time of the Posttest Plotted Against Angular Disparity and Condition for the Combined Data of Both Sexes.



Note. Control, C; stereotype threat, CS; mindfulness, MF; mindfulness and stereotype threat, MFS. Dots represent the mean. Bars represent the standard deviation. Both are slightly offset for the groups for better visibility.

For reaction time, model construction resulted in a model with random intercepts and slopes for PP and DEG by participant. PP*MF*ST*SEX*DEG and all respective interactions and main effects were analyzed as fixed effects. Significant differences were found for DEG*SEX*ST and DEG*PP (see Table 9). The interaction DEG*PP shows a significant increase in reaction time by DEG, with higher increases in pretest than posttest. Partial interactions show a significant decrease in reaction time for SEX*ST and DEG*ST. Reaction time increased significantly by DEG.

Table 9

Statistical Analysis of Reaction Time (in Seconds).

Variable	Estimate	SE	Test Statistic	*p* value	95% CI
Intercept	1.89	0.15			1.59, 2.21
Hypothesis-relevant Effects					
MF*ST*PP*SEX*DEG			$\chi^2(1)=0.01$	.927	
MF*ST*PP*SEX			$\chi^2(1)=2.20$	.138	
MF*ST*PP*DEG			$\chi^2(1)=1.14$	.286	
MF*ST*PP			$\chi^2(1)=0.09$	.771	
MF*PP			$\chi^2(1)=0.01$	.914	
ST*PP			$\chi^2(1)=2.39$	.122	
DEG*SEX			$\chi^2(1)=0.58$	.455	
DEG*PP(pre-post)	0.15	0.05	$\chi^2(1)=8.11$	.004	0.04, 0.24
DEG(0°)*PP(pre-post)	0.22	0.20			-0.17, 0.60
DEG*PP(post)	1.12	0.12			0.90, 1.36
DEG(100°)	1.17	0.06	$\chi^2(1)=164.10$	<.001	1.05, 1.28
Additional Effects					
DEG*SEX*ST			$\chi^2(1)=5.25$	.022	
DEG*(ST-noST)	-0.24	0.11	$\chi^2(1)=4.12$	.041	-0.49, -0.03
DEG*(noST)	1.21	0.09			1.04, 1.38
DEG(0°)*(ST-noST)	0.39	0.20			-0.03, 0.81
SEX(w-m)*(ST)	-0.55	0.27	$\chi^2(1)=3.94$	.047	-1.13, 0.00
SEX(w-m)*(noST)	0.21	0.19			-0.16, 0.59
SEX(m)*(ST)	0.29	0.19			-0.11, 0.67
MF*ST*DEG			$\chi^2(1)=3.05$	.081	
MF*PP*DEG			$\chi^2(1)=3.21$	.073	

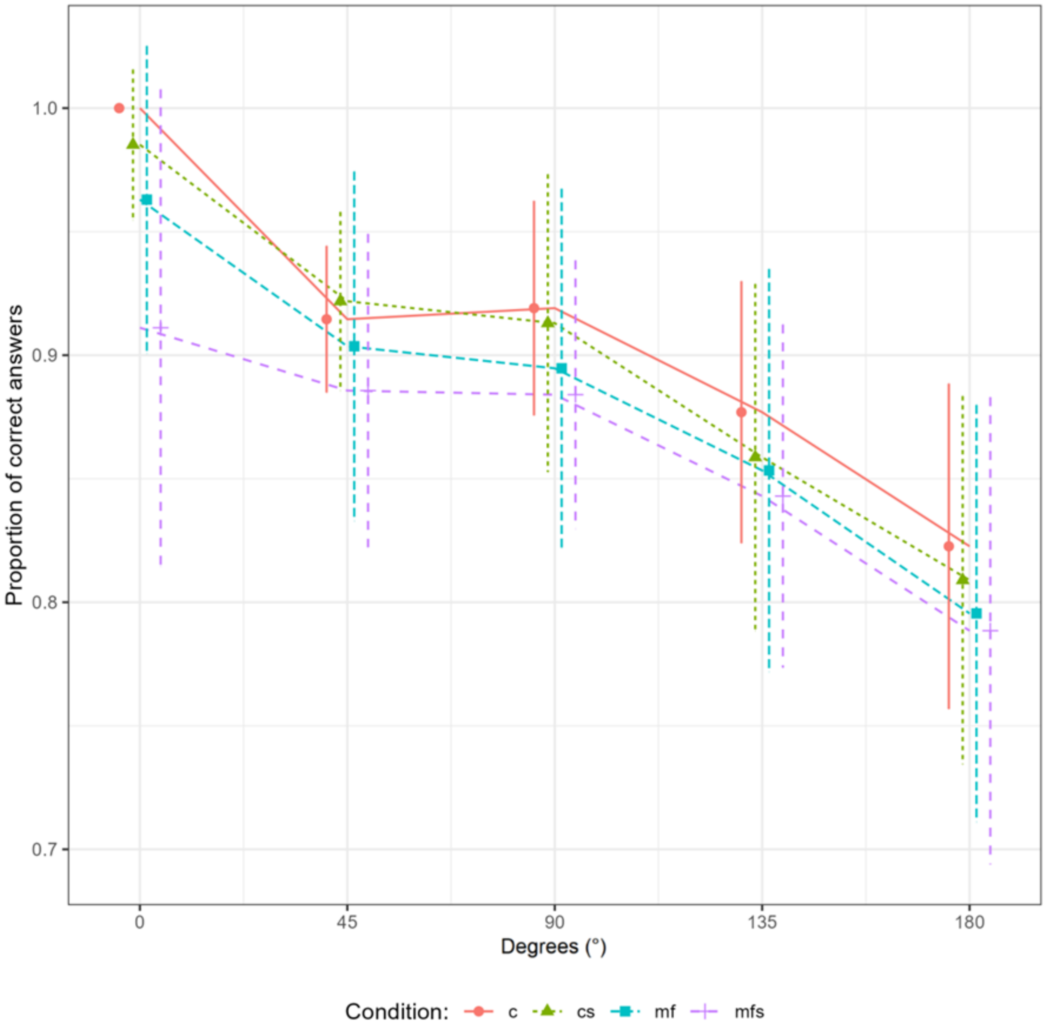
Note. Intercept in this model represents the estimate for SEX = men, no stereotype threat, angular disparity (DEG) = 0°, pre-posttest (PP) = post, and no mindfulness (no MF). Effect of SEX represents the difference between women (w) and men (m). Effect of ST represents the difference between stereotype threat (ST) and no stereotype threat (no ST). Effects of angular disparity (DEG) represent changes of 100°.

6.4.3 Accuracy

In Figure 16, the graphs of the proportion of correct answers are closely related. The visually interconnected points of the means show, that all graphs are slightly more apart at zero degrees (with large standard deviations; noticeably, the control group had only correct answers) and show a negative slope with increasing angular disparity.

Figure 16

Accuracy of the Posttest Plotted Against Angular Disparity and Condition for the Combined Data of Both Sexes.



Note. Control, C; stereotype threat, CS; mindfulness, MF; mindfulness and stereotype threat, MFS. Dots represent the mean. Bars represent the standard deviation. Both are slightly offset for the groups for better visibility.

Model construction for ACC resulted in a model with random intercepts and slopes for PP and DEG by participant. PP*MF*ST*SEX*DEG and all respective interactions and main effects were analyzed as fixed effects. Significant differences were only found for DEG (see Table 10). Accuracy decreased significantly by DEG.

Table 10

Statistical Analysis of (Logarithmic Odds of) Accuracy.

Variable	Estimate	SE	Test Statistic	*p* value	95% CI
Intercept	3.62	0.19			3.23, 4.06
Hypothesis-relevant Effects					
MF*ST*PP*SEX*DEG			$\chi^2(1)=0.01$	.918	
MF*ST*PP*SEX			$\chi^2(1)=1.00$	.317	
MF*ST*PP*DEG			$\chi^2(1)=2.31$	.129	
MF*ST*PP			$\chi^2(1)=0.03$	.871	
MF*PP			$\chi^2(1)=0.01$	.926	
ST*PP			$\chi^2(1)=1.21$	.272	
DEG*SEX			$\chi^2(1)=1.55$	.214	
DEG(100°)	-1.02	0.09	$\chi^2(1)=90.95$	<.001	-1.21, 0.83
Additional Effect					
MF*PP*DEG*SEX			$\chi^2(1)=3.63$	.057	

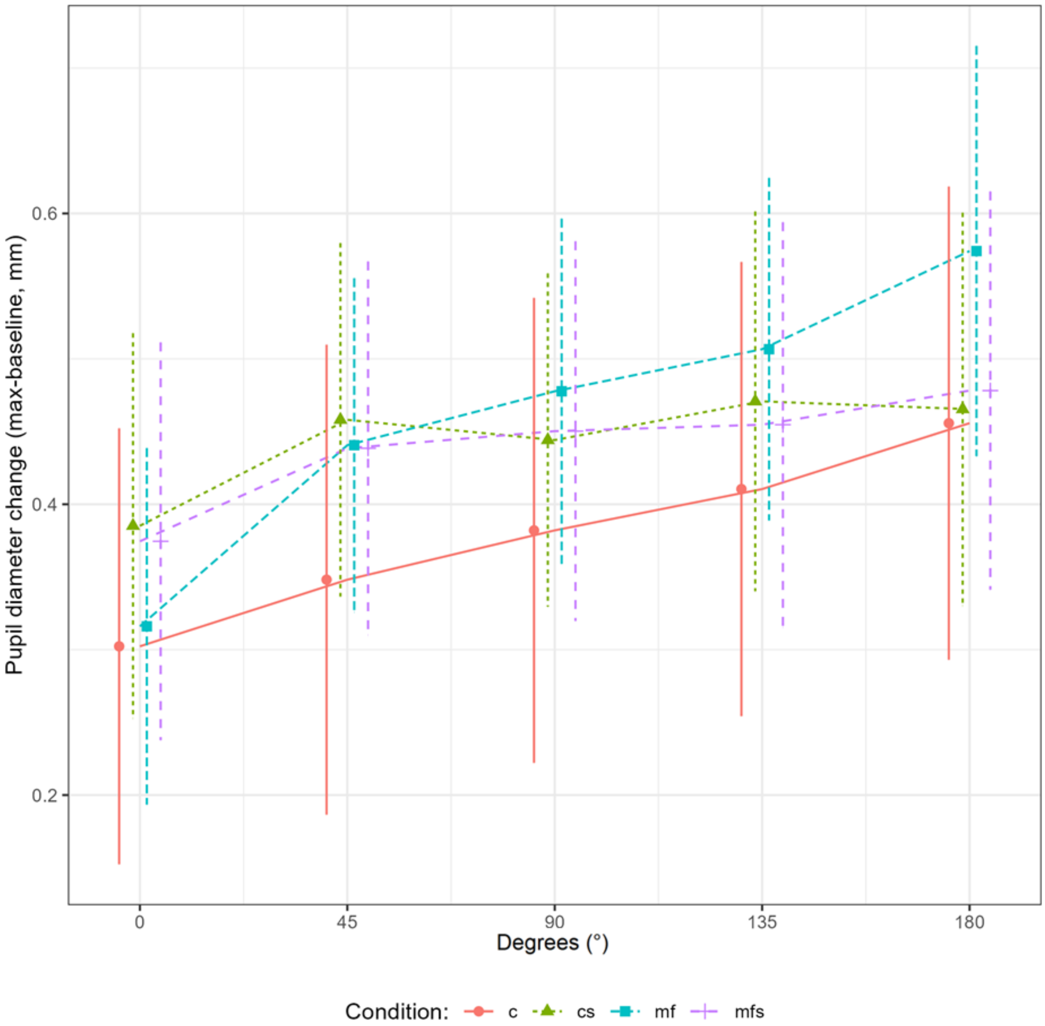
Note. Intercept in this model represents the estimate for the logarithmic odds at angular disparity (DEG) = 0° for SEX = men, no stereotype threat, pre-posttest (PP) = post, and no mindfulness (no MF). Effect of SEX represents the difference between women (w) and men (m). Effect of ST represents the difference between stereotype threat (ST) and no stereotype threat (no ST). Effects of angular disparity (DEG) represent changes of 100°.

6.4.4 Pupil Diameter

The task-evoked pupil responses are depicted in Figure 17 and show an inclination of all graphs with increasing angular disparity. Here, the control group shows constantly lower values than the other conditions. The graphs of the other conditions lie closer together.

Figure 17

Changes of Pupil Size (Max-Baseline) of the Posttest Plotted Against Angular Disparity and Condition for the Combined Data of Both Sexes.



Note. Control, C; stereotype threat, CS; mindfulness, MF; mindfulness and stereotype threat, MFS. Dots represent the mean. Bars represent the standard deviation. Both are slightly offset for the groups for better visibility.

The model building for the pupil diameter differences resulted in a model with random intercepts and slopes for PP and RT by participant. PP*MF*ST*SEX*DEG, RT, and all respective interactions and main effects were analyzed as fixed effects. The pupil diameter increased significantly by PP*SEX, DEG*ST, and RT (see Table 11). Pairwise comparisons showed significant differences between the groups for PP (woman/man), but not for SEX (pretest/posttest). The effect of RT in the model indicates an increase of the pupil size change with increasing reaction times.

Table 11

Statistical Analysis of Pupil Diameter (in 10^{-1} mm).

Variable	Estimate	SE	Test Statistic	*p* value	95% CI
Intercept	3.15	0.53			1.05, 5.00
Hypothesis-relevant Effects					
MF*ST*PP*SEX*DEG			$\chi^2(1)=0.06$	.812	
MF*ST*PP*SEX			$\chi^2(1)=0.30$	.587	
MF*ST*PP*DEG			$\chi^2(1)=0.30$	.585	
MF*ST*PP			$\chi^2(1)=0.31$	.580	
MF*PP			$\chi^2(1)=1.98$	.159	
ST*PP			$\chi^2(1)=2.15$	.143	
DEG*SEX			$\chi^2(1)=0.87$	.366	
RT	0.33	0.02	$\chi^2(1)=166.46$	<.001	0.30, 0.37
DEG(100°)	0.23	0.52	$\chi^2(1)=22.97$	<.001	0.14, 0.37
PP(pre-post)*SEX(w-m)	-0.88	0.43	$\chi^2(1)=3.94$	.047	-1.72, -0.04
PP(pre-post)*SEX(m)	3.11	0.38			2.34, 3.89
PP(post)*SEX(w-m)	-0.10	0.61			-1.22, 1.13
Split Interaction Main Effects					
PP	2.65	0.31	$\chi^2(1)=25.93$	<.001	2.05, 3.26
PP(w)	2.23	0.38	$\chi^2(1)=20.63$	<.001	1.50, 2.99
PP(m)	3.11	0.37	$\chi^2(1)=27.40$	<.001	2.34, 3.84
SEX	-0.74	0.52	$\chi^2(1)=1.89$	.169	-1.72, 0.26
SEX(pre)	-0.83	0.54	$\chi^2(1)=2.33$	.127	-1.93, 0.23
SEX(post)	-0.13	0.60	$\chi^2(1)=0.05$	.827	-1.26, 1.07
Additional Effects					
DEG*(ST-noST)	-0.28	0.09	$\chi^2(1)=9.29$	.002	-0.46, -0.09
DEG*(noST)	0.38	0.07			0.23, 0.51
DEG(0°)*(ST-noST)	0.15	0.52			-0.92, 1.25
DEG*ST*PP			$\chi^2(1)=2.97$	.085	

Note. Intercept in this model represents the estimate for SEX = men, no stereotype threat, angular disparity (DEG) = 0°, pre-posttest (PP) = post, and no mindfulness (no MF). Effect of SEX represents the difference between women (w) and men (m). Effect of ST represents the difference between stereotype threat (ST) and no stereotype threat (no ST). Effects of angular disparity (DEG) represent changes of 100°.

6.5 Discussion

6.5.1 Toronto Mindfulness Scale Scores

All participants showed overly high scores for state mindfulness in the experiment (see Lau et al., 2006). The scores are similar to the ones Weger et al. (2012) reported for their mindfulness groups. Interestingly, even the participants without mindfulness intervention in this study showed high levels of state mindfulness in the scores, in contrast to significantly lower scores of the respective group in Weger et al. (2012). It is important to note that the Toronto Mindfulness Scale scores indicate that the brief mindfulness intervention did not have a significant main effect on state mindfulness, which was not the expected effect to form the basis for the rest of the analysis. Thus, this must be considered regarding the following interpretations of the results.

6.5.2 Mental Rotation Performance

Based on our hypothesis, the focus was on the interaction between mindfulness, stereotype threat, and pre-post testing, which would illustrate the changes resulting from both treatments. We expected these changes to be larger for tasks of higher difficulty, which is why angular disparity was added to the modeling. Additionally, we examined exploratively whether sex had an influence.

No significant interactions between pre-post testing, mindfulness, and stereotype threat emerged in the analysis for reaction time and accuracy, which indicates that no effects regarding the treatments of stereotype threat and mindfulness can be observed. For accuracy, the resulting model is even more reduced than the one for reaction time, with angular disparity being the only significant variable. A possible explanation is that in mental rotation tasks, a speed-accuracy-tradeoff always happens. Here, accuracy undergoes smaller changes than reaction time and is therefore less influenced by different variables (see e.g., Hertzog et al., 1993; Wickelgren, 1977).

In general, all groups improved in their reaction times from pretest to posttest. However, no significant differences in the performance changes manifested between the four condition groups. This statistically renders both treatments, that is, mindfulness and stereotype threat, and their respective combinations, as overall ineffective concerning all dependent variables in this study. Additionally, the results show no significant sex differences in overall behavioral performance. The main effects and interactions for both accuracy and reaction time do not indicate any influence of sex in the models. This is in line with other chronometric mental rotation studies. For instance, Voyer et al. (2006), and Jansen-Osmann and Heil (2007) also reported no sex differences in mental rotation performance of 3D cube figures.

Altogether, our results did not confirm our hypothesis that task performance would be better after the mindfulness induction in the stereotype threat condition. Additionally, all results apply to both sexes in the same way in this study. This stands in contrast to the findings of Weger et al. (2012), on which our hypothesis was based. With an applied stereotype threat, the authors reported better performance due to a brief mindfulness intervention. With our intervention being recorded by a professional instructor and lasting the same amount of time, that is, five minutes, the intervention aspects of both studies can be expected to be very similar and not likely the cause for the differences.

First of all, the missing stereotype threat effect is astonishing, because in general, laboratory effects of stereotype threats are robust, with moderate to small effect sizes (for a review, see Spencer et al., 2016). Nguyen and Ryan (2008) reported subtle cues to trigger larger stereotype threat effects for women in math than did blatant or moderate ones. However, in order to remove threat, blatant strategies reduced stereotype threat effects sizes more than subtle strategies did. In this study, the stereotype threat cue could have been too unconcealed to elicit a sufficiently strong effect. Additionally, the mindfulness intervention was not targeted on removing the stereotype threat per se, and cannot be described as a blatant strategy to reduce it.

However, its effects were expected to reduce stereotype threat effects as a consequence. The findings of Nguyen and Ryan (2008) also indicated that stereotype threat effects for women in math were largest among women who moderately identified with math. Spencer et al. (2016) described this as an unfortunate aspect of stereotype threat: that the most dedicated and caring people are affected most by negative stereotypes. Participants' performance diminishes more under stereotype threat, when they identify with the stereotype domain, with their performance in the domain being self-relevant (Spencer et al., 2016). Here, mental rotation tasks might be different from math tasks. At school, people are confronted with math for several years and establish an attitude toward the domain and their respective abilities in it. Specifically, women's performance can suffer through stereotype threat, because they are "aware of the widely held stereotype threat in our Western culture that women are not as good at math as men" (Doyle & Voyer, 2016, p. 10; see also Matlin, 2011; Nosek et al., 2002; Tartre & Fennema, 1995). The participation in a mental rotation experiment might not cause similarly strong effects, because people would not be able to identify and range their own performance level in this domain right away. Even more, after the practice trials, no more indication for correct or wrong answers was given in this experiment. The participants were therefore oblivious to how well they were currently performing, in contrast to mathematical tasks, where the own competence might be evaluated to a greater extent. Hence, low levels of stereotype threat effects regarding mental rotation task performance might be due to the unfamiliarity with the test and the lack of knowledge about one's own competence compared to others. Here, the difference between psychometric and chronometric mental rotation experiments might also have an effect. Sex differences have been reported more in psychometric experiments than in chronometric studies (Voyer et al., 1995). It is possible, that the chronometric tests do not only show smaller sex differences in performance but are also perceived as less stereotypical. However, it is generally the mental rotation ability and not the test design that is assumed to be stereotypical. For example, Hausmann et al. (2009) activated and found gender stereotypes for the item “imaging abstract objects and rotate them mentally”, which should apply to both psychometric and chronometric tests. However, the activated stereotypes only resulted in performances difference in a psychometric test and not in other spatial abilities (with no direct comparison with a chronometric test). Thus, while it is possible that stereotypes exist for spatial abilities in general, these might transfer to performance differences only in specific tests. Furthermore, it has been shown that not all women suffer from stereotype threat activation in mental rotation. In one study, this stereotype threat effect was only visible in women with a feminine gender role orientation (Tempel & Neumann, 2016). For future research, it might be worth analyzing the gender role orientation. It is also possible that task difficulty was not sufficiently high enough. Other studies have shown that only very difficult tasks were affected by stereotype threat (Spencer et al., 1999; Blascovich et al., 2001; Keller, 2007).

Another possible explanation for our interventions not to show any effects might be our test sample. Facing stereotype threat, strong coping abilities are associated with resilience. Inzlicht and colleagues (Inzlicht, Aronson, et al., 2006; Inzlicht, McKay, et al., 2006) reported high self-monitors not to underperform under stereotype threat. The same applies to people with a high coping sense of humor (Ford et al., 2004). The participants in this study were students of sports and movement sciences. It is possible that this encompasses better coping mechanisms, that is, showing stronger minds in terms of failure (Seeley & Gardner, 2003) and therefore not underperforming under stereotype threat (Inzlicht, Aronson, et al., 2006; Inzlicht, McKay, et al., 2006). The results of the analysis of the Toronto Mindfulness Scale scores are in line with these interpretations (see above). The high TMS scores might indicate one reason, why the mindfulness treatment was statistically insignificant. Having all test groups on high scores already, the extent, to which an intervention might elevate state mindfulness, might get rendered insignificant.

In line with the paradigm for chronometric mental rotation tasks, changes in angular disparity significantly influenced all dependent variables. Higher angular disparity between the two pictures resulted in higher reaction times and lower accuracy. Differences between the reaction times of pre- and posttest can be observed, with higher values in the pretest. The interaction between pre-post and angular disparity shows a higher reaction time increase for each degree in the pretest. In addition, the performance increase from pre- to posttest is higher at greater angles. These outcomes can be explained as typical learning effects in a pre-post experimental design, and have more impact for more difficult tasks, that is, at higher angular disparities.

6.5.3 The Role of Cognitive Effort in the Mental Rotation Task

Before addressing the hypotheses, we would like to elaborate on possible problems in the analysis of pupil size means. In the first study of this thesis (Bauer et al., 2022), we exploratively conducted both a baseline correction similar to this study and a trial-dependent baseline correction (see Mathôt, 2013; Mathôt et al., 2018; Mathôt et al., 2015), which lead to the same results. Both types of baseline correction have advantages (see Holmqvist et al., 2011; Mathôt et al., 2018). However, by using short breaks between the trials (i.e., fixation point sections), the risk of possible carry-over effects should be of concern for mental rotation experiments. Considering our fixation point duration of 500 ms in this study, we used the described baseline correction, as it should be slightly better regarding possible carry-over effects. In our last paper, we pointed out that future experimenters should 1) keep showing the task and measuring pupil dilation even after the response for at least 500 ms, and 2) increase the break between trials to at least two seconds (see Bauer et al., 2022).

Regarding the analysis of pupil diameter differences as a measure for cognitive load, we expected reduced load for the mindfulness and stereotype threat condition. The significant interaction of sex and pre-post testing indicates that a cognitive improvement in form of lower cognitive load was higher for men than for women, which might be due to men having started at a higher level and having more room for improvement than women had (i.e., a ceiling effect affecting the women more). These differences were not significant in the main effect for sex, however. The main effect for pre-post testing shows an improvement of both sexes over the two experiment parts. Furthermore, the interaction solely shows a sex difference for the temporal effects of pre-post testing, for instance, through greater learning improvements in men, but does not seem to interact with any of the treatments of the experiment. Neither mindfulness nor stereotype threat had a statistically significant impact on the cognitive load for both sexes. This finding is congruent with our results for behavioral performance.

Applied stereotype threat leads to the motivation in affected people to disconfirm negative stereotypes that target their social identity or at least to avoid confirming it (for a review, see Spencer et al., 2016). This inflicts a pressure to succeed that can negatively influence the performance through three mechanisms: mere effort, working memory depletion, and conscious attention to automatic processes (Spencer et al., 2016). Based on this, stereotype threat should indeed increase cognitive load in laboratory experiments, which we also predicted in our hypothesis. As discussed above, our stereotype threat application as well as our test sample with already high state mindfulness might not have been ideal to establish an impactful stereotype threat on the participants. However, for this study, our third hypothesis was rejected.

In the model, changes in pupil dilation did not get fully explained by variations in angular disparity. Reaction time itself still predicted a significant portion of the pupil diameter in the statistic model, which is in line with the findings of the first study (Bauer et al., 2022). That is, both reaction time and angular disparity had an impact on cognitive load, but none of them alone seemed to be sufficient to describe the connection between each other, and to account for task difficulty. The relationship between task performance and cognitive load is also modulated by task difficulty, which is hard to control (Coyne et al., 2017). Therefore, the relationship between difficulty and cognitive load is probably not linear (Bauer et al., 2022).

In the study of Campbell et al. (2018), women showed higher cognitive load—or used more attentional resources—than men in mental rotation tasks of abstract figures. Our findings did not confirm this, as no differences emerged. As a conclusion, both sexes were able to solve mental rotations tasks with cube figures of similar difficulty with the same cognitive effort. The first study showed similar results regarding the sex differences in cognitive effort (Bauer et al., 2022). Hence, our findings in this study provide further evidence that men and women need the same cognitive effort for the solution of cube figure tasks. The possible effects might get analyzed further in future studies using different kinds of stimuli and stereotype threat application to be further illustrated.

6.6 Limitations

In pupillometry studies, the pupil foreshortening error should be of concern to have an influence on the acquired data, especially when looking at areas, which are further away from the screen center (Hayes & Petrov, 2016). In our mental rotation tasks, all analyzed image pairs are shown at the same positions on the screen. Possible pupil foreshortening errors were not expected to systematically differ between conditions, and were neglectable for that reason (see Mathôt et al., 2018).

We conducted this experiment with close similarity to the study of Weger et al. (2012). They presented the stereotype threat in one sentence in the instructions and did not include a manipulation check after that. We replicated their approach, taking also into account the already long participation time for the experiment. Retrospectively, we would recommend using a manipulation check questionnaire, as, for instance, Neuburger et al. (2015), and Pennington et al. (2019) did. It is also important to note that the data for this study was acquired in only one country. Gender stereotypes differ across countries, which may influence the performance in gendered abilities (Moè et al., 2021). Another consideration could be that gender stereotypes within age groups may change over the years (Bhatia & Bhatia, 2021).

One more limitation is the short mindfulness induction. As it was the objective to research the effects of a single bout if brief mindfulness induction, it is still important to consider that mindfulness general works best through daily practice routine over a longer period of time (e.g., Rahe & Jansen, 2023).

As mentioned in the discussion, the experimental design led to possible carry-over effects of cognitive effort, which must be taken into account for interpretation of the results.

6.7 Conclusion

No sex differences appeared in either of the measurements, which contributed to the different results in chronometric mental rotation test concerning sex differences (see Jansen-Osmann & Heil, 2007). This study replicated the experimental design of Weger et al. (2012) to scrutinize the impact of stereotype threat and mindfulness and showed that the here induced stereotype threat had no impact on mental rotation task performance, which was successfully applied in former mental rotation studies, but with psychometric tests (Heil et al., 2012). However, we found that high levels of state mindfulness inhibit stereotype threat to negatively influence task performance and cognitive load. This must be investigated in more detail.

7 General Discussion

The aim of this dissertation was to gather insights about the effects of embodiment, mindfulness, and stereotype threat on chronometric mental rotation task performance and cognitive load. Possible sex differences were of additional interest. In this chapter, I summarize the findings of the three studies, reflect and discuss how they contribute to the broader literature, and propose future directions. The chapter ends with a conclusion regarding all elements of this thesis (see Lewis et al., 2021).

7.1 Summary of Dissertation Findings

7.1.1 Objective 1: Mental Rotation with Abstract and Embodied Figures

In the first study, we investigated the behavioral performance and cognitive load in chronometric mental rotation tests with one abstract and two embodied stimuli types that were comparable in shape and color. Sex differences and embodiment effects were analyzed for the dependent variables reaction time, accuracy and cognitive load determined via pupillometry. Statistical analysis showed no sex differences for any of these dependent variables. However, the results indicated better behavioral performance and lower cognitive load for the embodied figures, as well as higher cognitive load for higher angular disparity between the figures. In the model, reaction time itself predicted a significant portion of the pupil size changes.

7.1.2 Objective 2: Mental Rotation with Embodied Meditation

The second study comprises the analysis of effects of an embodied (mindfulness) meditation intervention on chronometric mental rotation test performance and subjective cognitive load. Sex differences and embodiment effects were analyzed for the dependent variables reaction time, accuracy and self-reported cognitive load determined via a Likert-type scale after each test block. The main effects of the mindfulness induction were not significant in any of the dependent variables; however, a significant three-way interaction emerged between pre-post testing, mindfulness, and sex for reaction times (see Figure 11). Additionally, men's behavioral performance was more affected by increasing levels of angular disparity than women's. No sex differences emerged for subjective cognitive load.

7.1.3 Objective 3: Mental Rotation with Embodied Meditation and Stereotype Threat

The third study involved a brief mindfulness intervention and a stereotype threat induction. Their effects as well as sex differences were analyzed for chronometric mental rotation tests, with the dependent variables reaction time, accuracy and cognitive load determined via pupillometry. Statistical analysis showed no sex differences for any of these dependent variables. Furthermore, neither stereotype threat nor mindfulness influenced task performance significantly. However, we found that high levels of state mindfulness inhibit stereotype threat to negatively influence task performance and cognitive load.

7.2 Findings in the Context of Evolving Literature

7.2.1 Sex Differences in Chronometric Mental Rotation Task Performance

The possibility of sex differences in mental rotation is an important topic of cognitive psychology research, as the found differences favoring men are substantial in psychometric mental rotation tests (e.g., Voyer et al., 1995). The interest in the analysis of sex differences remains high, as the performance in both mental rotation and STEM fields is connected and might also be an important indicator for the presence of fewer women in the STEM fields (see Law et al., 2021; Morais Maceira, 2017; Makarova et al., 2019). For chronometric mental rotation tests, research shows smaller and mostly non-significant differences (e.g., Jansen-Osmann & Heil, 2007; Peters & Battista, 2008). Therefore, researchers examine and theorize on the nature of and differences between psychometric and chronometric mental rotation tests, for instance, to what extent the mental rotation ability or other aspects of the whole cognitive task are involved and lead to the found differences between men's and women's performance (e.g., see Jost, 2022). Additionally, with the development of computer-based analysis, both forms of mental rotation tests are progressively investigated further.

Concerning the chronometric mental rotation tests, the three studies of this thesis add to the broad findings in the literature, as none of them show significant sex differences regarding the behavioral performance. This is indicated by the respective interactions and main effects in all studies. Only in the second study, the reaction time analysis showed a significant three-way interaction of pre-post testing, the mindfulness treatment, and sex, which could indicate a difference on how men's and women's performance changed due to the mindfulness condition. As mentioned in the second study, splitting up this interaction showed no further significant results, which could be analyzed with more test power in future studies. Overall, one main finding of this thesis is that the embodiment effects elicited by test stimuli as well as one-time embodied mindfulness interventions seem to affect men and women similarly in their chronometric mental rotation performance.

As for the analyses of cognitive load, no significant sex differences emerged in any of the three studies. The first and third study included pupillometry as a measure of changes in cognitive load per item, and the second study included a Likert-type scale to ascertain subjective cognitive load for whole test blocks. To date, there is not yet much literature regarding sex differences in cognitive load in mental rotation tests. In contrast to the pupillometric findings of Campbell et al. (2018) and the findings for subjective cognitive effort of Jost et al. (2023), our results showed similar performance in the mental rotation tasks and similar levels of cognitive effort for each stimulus type for both genders. These results regarding embodiment effects of the stimulus material coincide with the findings of the second and third study, which neither showed sex differences for embodiment effects via focused attention meditation (mindfulness) nor for the interaction with stereotype threat. The results for cognitive load are like those for behavioral performance, which can be explained by both mental rotation performance (e.g., Jost & Jansen, 2020) and cognitive load (e.g., Kahneman & Beatty, 1966) being linked to task difficulty (see **Section 7.2.3**).

In summary, all three studies of this thesis provide evidence that men and women use similar cognitive effort to solve chronometric mental rotation tasks.

7.2.2 Embodiment in Chronometric Mental Rotation Tests

According to Embodied Cognition, the brain incorporates experiences of the whole body and combines sensory and motoric information (Barsalou, 2008; Lakoff & Johnson; 1999). Hence, embodiment effects can influence the performance in cognitive tasks. Some of these effects are observed in embodied mindfulness meditation, which can improve attentional control capacity and redirect attention to a current task (Chiesa et al., 2011).

Regarding the effectiveness of brief mindfulness interventions, Zeidan, Johnson, et al. (2010) found that four days of twenty-minute mindfulness meditation training seemed to be able to increase the capability to sustain attention. Comparing a mindfulness group and a control group—both without prior meditative experience and no difference at baseline measures—they found significant effects on cognitive tasks that require sustained attention and executive processing efficiency (Symbol Digit Modalities Test, verbal fluency, and hit runs on n-back task). The brief mindfulness training led to reduced fatigue and anxiety ratings compared to the control group. Zeidan, Johnson, et al. (2010) further concluded that fatigue and anxiety may specifically have an impact on information processing. Mindfulness emphasizes achieving a balance between a vigilant and relaxed state of mind (Wallace, 2006); and being able to self-regulate emotions was indicated to be essential for enhancing cognition (Austin, 1998; Moore & Malinowski, 2009; see Zeidan, Johnson, et al, 2010). Consequently, the mindfulness process of increased present moment awareness combined with returning attention to a specific object can boost attentional stability (Epel et al., 2009; Wallace, 2006; see Zeidan, Johnson, et al, 2010). Brief mindfulness training may not only improve mood, but also increase cognitive processing skills, such as reducing lapses of attention. One short-term beneficial effect might be an increased focus in timed or speeded tasks (Zeidan, Johnson, et al, 2010). Similarly, Tang et al. (2007) reported five days of Integrative Body Mind Training having improved cognitive processes and mood. However, as this training includes various techniques (e.g., music therapy, mindfulness, and guided-imagery), it is unclear how much of an impact mindfulness had for the improvements. Zeidan, Johnson, et al. (2010) point out that extensive training is definitely more effective than brief mental training, as consistent training develops persisting changes in well-being and cognition (Cahn & Polich, 2006). However, they conclude that shorter training offers extensive benefits by being more attractive and easily disseminated, which might persuade individuals to start and continue meditation practice, having experienced an effect after one brief training (Grossman et al., 2004; Zeidan, Johnson, et al, 2010).

Regarding even shorter and one-time interventions, in a recent study, Bokk and Forster (2022) analyzed the effects of a ten-minute mindfulness intervention on the P300 event-related potential, which is a neurophysiological marker of attention (Polich, 2012) and known to be highly dependent on expectancy (Bokk & Forster, 2022). They employed a classical oddball paradigm, where the somatosensory P300 decreases with higher block numbers (Kida et al., 2012), which indicates the task not being difficult enough, leading to mind-wandering (Picton, 1992; see Bokk & Forster, 2022). In meditation-naïve participants, the mindfulness intervention prevented the decrease of somatosensory P300 normally observed with task repetition (Datta et al., 2007; Kida et al., 2012; Lammers & Badia, 1989; Nakata et al., 2015; see Bokk & Forster, 2022), which was observed in their control group. This decrease with task repetition suggests a habituation effect of diverting attention away from the task (Isreal et al., 1980; Wickens et al., 1983). The mindfulness group not showing such an effect may suggest either changes in attention mechanisms or a preservation of these mechanisms against habituation effects (Bokk & Forster, 2022). The authors concluded that "even a short mindfulness meditation prevents the depletion of attentional resources on the task" (Bokk & Forster, 2022, p. 2027).

In this thesis, the only significant embodiment effects emerged in the first study, where the results showed better behavioral performance and lower cognitive load for the embodied figures. The second and third study involved the analysis of embodiment effects through a single focused attention meditation, one form of mindfulness meditation. Neither the twenty-minute body-scan nor the five-minute raisin-eating intervention led to significant main effects of the mindfulness condition.

Considering the respective limitations of the studies, overall, the thesis's results provide evidence that a single bout of brief mindfulness intervention does not give rise to significant embodiment effects, neither in mental rotation performance nor in cognitive load. Therefore, to elicit a significant improvement in the cognitive task and to reduce cognitive load, both brief interventions did not sufficiently influence the introspective systems of the brain, which include consciously experiencing emotion and cognitive operations (see Niedenthal et al., 2005). These findings stand in contrast to other studies, which reported significant effects, for instance, in mathematics (Weger et al., 2012) or psychometric tests (Rahe & Jansen, 2023).

However, to our knowledge, the second and the third study of this thesis are the only ones to date to analyze these effects in chronometric mental rotation, and with the addition of cognitive load. Therefore, these studies are a small contribution to fill the gap of evidence about whether brief one-time mindfulness intervention could already induce significant changes in the cognitive performance of chronometric mental rotation tasks and respective cognitive load. As our findings suggest, more instances than only one short mindfulness induction seem to be needed to elicit significant effects related to embodied cognition and cognitive load. While the elevated state mindfulness scores indicate effective induction in the treatment groups, the embodiment effects were statistically not different to the control groups.

7.2.3 Cognitive Load and Mental Rotation Task Difficulty

An interesting aspect of the analyses of cognitive load in all three studies is that in all models—both for objective and subjective measurements of cognitive load—the variable reaction time (with a slight variation in the second study) is always a significant fixed effect. As both task performance and cognitive load are dependent on and modulated by task difficulty (Coyne et al., 2017; Jost & Jansen, 2020; Kahneman & Beatty, 1966), the consistent findings of all three studies indicate that the relationship between task difficulty and cognitive load might not be linear. Thus, while showing trends like behavioral performance, cognitive load might behave differently compared to the general paradigm of mental rotation, according to which performance is linearly dependent on task difficulty, which is to some extend due to higher angular disparity (e.g., Shepard & Metzler, 1971). Regarding the question of a linear dependency of cognitive load on task difficulty, in studies on updating tasks, in comparison, overall pupil dilation also increases with higher task demands. However, when task demands exceed the available cognitive resources, pupil dilation does not seem to increase anymore (see van der Wel & van Steenbergen, 2018). For instance, Poock (1973) found that when the cognitive capacity is under overload, the average pupil dilation falls below baseline. Other studies using a digit span test indicated pupil dilation to occur until reaching an asymptote at a string of nine (Granholm et al., 1996) or thirteen digits (Peavler, 1974), and pupil constriction to occur beyond that. The timing of the latter seems to be linked to the instructions, that is, whether they point out to maintain active rehearsal under overload (Granholm et al., 1996; see van der Wel & van Steenbergen, 2018).

In summary, pupil dilation increases with higher task difficulty, but when working memory reaches capacity limitations, the pupil dilation can plateau or drop, which may be modified by task instructions (van der Wel & van Steenbergen, 2018). Additionally, the relationship between behavioral performance and cognitive effort can also only be reliably investigated, when neither floor nor ceiling effects are present in both variables (Norman & Bobrow, 1975; see Hockey, 1997; van der Wel & van Steenbergen, 2018). From the viewpoint of motivation, very high task demands may lead to disengagement and withdrawal of effort, which describes an inverted U-shaped relationship between task difficulty and effort (Brehm & Self, 1989; also see van der Wel & van Steenbergen, 2018).

In mental rotation, task difficulty can be influenced and determined by stimulus complexity and angular disparity in different rotation axes. Depending on these factors, both floor and ceiling effects have always to be considered to possibly impact the analysis. While the graphs of the combined data of pupil diameter differences in the first (Figure 5) and third study (Figure 17) generally follow a trend of linear increase, on visual inspection, the increase in the third study could be interpreted to be slightly reduced and plateau at higher angular disparity for the conditions including stereotype threat induction. This could mean that the stereotype threat induction restricted working memory capacity. As such, this tentative interpretation would be in line with the findings of the studies on updating tasks (see van der Wel & van Steenbergen, 2018) as well as our results of the interaction between angular disparity and stereotype threat in the statistical analysis. However, in mental rotation, the findings are still unclear, and future research should analyze this relationship between cognitive effort, task performance, and task difficulty in more detail.

7.2.4 Mindfulness and Stereotype Threat

In mental rotation studies, abstract cube figures are frequently used as stimuli. They can be regarded as more male stereotyped and considered as one reason for possible performance differences (Rahe et al., 2021; Rahe & Quaiser-Pohl, 2020; Ruthsatz et al., 2014, 2015). While this may be true on a population level, our results showed no sex differences in the first and second study, and especially in the third study, where the nature of the figures might have had an effect with the stereotype induction, such as an enhanced stereotype threat for women.

In their recent study with younger and older adolescents solving a psychometric mental rotation test, Rahe and Jansen (2023) reported a non-significant effect of stereotype activation on behavioral performance. The findings of our third study are in line with their results, but not with the results on mathematical performance of Weger et al. (2012), who did find significant effects with an auditory induction. Overall, the second and the third study confirm the existing literature that mindfulness practice can be a useful tool to elevate state mindfulness. While the third study indicated that mindfulness can inhibit stereotype threat to negatively influence task performance and cognitive load, the interaction with stereotype threat has to be further analyzed in the future.

7.3 Limitations

The three mental rotation studies with analyses of embodiment effects and cognitive load resulted in a slightly wider scope on the topic, as the first study involved effects within the test, and the others dealt with treatment effects. Also, due to the COVID-19 pandemic, the second study could only be designed as an online experiment, which entailed a different measurement of cognitive load being used. However, this thesis shows the similarities and interconnectedness of the analyses of embodiment effects, cognitive load, and sex differences, and with the third study also including stereotype threat. To some extent, the thesis points out the current limited knowledge of mental rotation and emphasizes the necessity for further research in this field.

In all the studies of this thesis, the participants shared mostly similar demographic characteristics such as affiliation with the same university or study program, age range, and German mother tongue. Some participants may have been involved in more than one study. Furthermore, at the institute, several mental rotation experiments are being conducted each year. To prevent more recent learning effects, participating students were required not to have taken part in any mental rotation study at least six months prior to each study of this thesis. In addition, in the third study, participants were required not to have any previous practical experience with mindfulness. Still, a transfer to the general population might not be entirely clear, which is a common problem in studies with student samples (e.g., see Hanel & Vione, 2016). Additionally, all participants in the studies were cisgender. As we investigated sex and/or gender differences in behavioral performance and cognitive load within this scope, it was impossible to distinguish between sex and gender in the analyses.

Although we endeavored to conduct all experiments, data processing, and analyses in accordance with state-of-the-art standards and recommendations, future advancements might offer new insights and better suited approaches than the currently used methods. Here, regarding the recent discussion on the replication crisis in psychology, one limitation may be that we had not pre-registered the three studies; however, we conducted all studies in the same manner, as if we had. Finally, because of my personal biases and constraints, this thesis may lack sufficient depth in some areas to appropriately acknowledge and highlight specific aspects, which may also include the uncertainty of the proposed conclusions. Because of that, all interpretations and implications should be considered cautiously and further scrutinized with future methods and information.

7.4 Implications for Education, Sports Practice, and Research

Although the results of this thesis mostly provide evidence for non-significant effects of one brief meditation intervention on chronometric mental rotation performance and respective cognitive load, I would like to elaborate on possible benefits and applications in practical fields, which may result from further research. As mindfulness would be exercised and used on a more regular basis in practical application, future research may investigate and elaborate, how many short sessions are needed for a beginner meditator to achieve consistent beneficial effects in chronometric mental rotation as well as other cognitive tasks and areas, and how short these sessions can be. As such, our findings may help to better understand the limitations of embodiment and mindfulness, while also considering the potential and benefits shown in other studies.

Overall, the studies of this thesis may add potential for further research, which could expand and deepen our knowledge and connect our findings to new insights. These findings imply possible transfers and generalizations of practical relevance for education and sports practice, which will be elaborated on in the following.

7.4.1 Education

Mindfulness exercises incorporate strategies that can simultaneously regulate attention and emotion, with various objectives such as the cultivation of well-being and emotional balance (Lutz et al., 2007, 2008). Mindfulness is normally taught and practiced over multiple sessions and then cultivated into a habit (e.g., MBSR). Normally, meditation beginners start with focused attention meditation and may expand their practices to open monitoring meditation later (see Lutz et al., 2008). Consequently, they can become more proficient in regulating their attention, which can lead to easier elevating their state mindfulness and maintaining these levels. Over time, focused attention meditation can induce a trait change of the practitioners to direct and stabilize their focus (Lutz et al., 2008).

These trait changes could become an important aspect in education, from elementary school (e.g., Portele & Jansen, 2023) and middle school pupils (e.g., McKeering & Hwang, 2019) up to university students (e.g., Bamber & Schneider, 2016). One positive effect could be the general training and resulting improvement of various mechanisms, which could lead to better learning and studying conditions. These mechanisms include directing and maintaining (cognitive) attention on specific objects and topics, perspective-changing capabilities, emotional regulation, and body awareness (see Hölzel et al., 2011). The increase in attentional control could lead to better grades, especially in children and adolescents who might have problems in staying focused. The introduction to such focused attention skills in a school setting could therefore be a good way to educate pupils to become better learners. In education, a lot of the already created programs—regarding mindfulness, mindfulness-based awareness, social and emotional learning, and others—have the purpose in common to support "children in cultivating a sense of well-being so they can thrive and flourish in the daily hustle" (Portele & Jansen, 2023, p. 2; see also Schonert-Reichl et al., 2015).

Regarding the emotion regulation through mindfulness (e.g., Amundsen et al., 2020; Portele & Jansen, 2023; Roemer et al., 2015), another benefit could be the mitigation of negative effects—for instance, stereotypes, stereotype threats, and anxiety—related to specific topics, such as mathematics, physics, or chemistry. Growing up with this support, less gender-stereotyped education, and respectively developed mindset, younger generations could be less influenced by such negative effects and have more interest and better grades in STEM topics and subjects related to mental rotation (see Rahe & Quaiser-Pohl, 2023). At the university, focused attention meditation courses could also be offered to students, in a similar way like scientific-writing or statistics classes, for example, through program orientations and onboarding courses (introducing concepts of mindfulness and services available on campus), or extra-curricular activities like tutoring and buddy programs (Bamber & Schneider, 2016; Martin et al., 2022). For example, mindfulness was incorporated at medical schools and offered an effective coping strategy for students (Hassed, 2021; Phillips, 2015). Mindfulness could easily be integrated into courses by doing a short meditation session at the start of didactic courses (Bamber & Schneider, 2016). If the education system and, consequently, younger generations incorporate these practices, mindfulness might increase success, decrease attrition, reduce stereotype impact, and improve academic achievement, for instance, through increased attentional control and emotion regulation such as reduction of anxiety and stress (see Bamber & Schneider, 2016). This could also lead to a more equal distribution of all genders in the STEM fields in the future. Additionally, it could reduce the gender gap and eliminate gender inequalities (e.g., see Breda et al., 2020; Campaña et al., 2023). The gender equality and equal opportunities on the labor market could also benefit the productivity and capacity of the economy, for instance, of the European Union (Morais Maceira, 2017).

7.4.2 Sports Practice

In addition to the benefits in education and their further reaching implications, there are more practical needs and advantages of progressing mindfulness research. Considering all the possible effects, the benefits can be pursued not only for cognitive tasks, but also for motoric tasks, which also involve high levels of, for instance, attentional control (see Ducrocq et al., 2016). As such, combined with emotion regulation, mindfulness can also be beneficial in high level sports. For instance, a five-week mindfulness program was found to boost the endurance performance and multiple cognitive functions of university athletes (Nien et al., 2020).

In another study, a mindfulness program of one weekly session of Flow Meditation for ten weeks reduced several variables regarding mood, impulsivity, and pre-competition anxiety-state. The findings indicated that mindfulness can increase the ability of athletes to cope with high-level sport pressure and to manage daily life better (Sánchez-Sánchez et al., 2023). Furthermore, Dehghani et al. (2018) compared a Mindfulness-Acceptance-Commitment (MAC) program with a control group by testing 31 students. The authors found that the MAC significantly increased the performance of basketball playing athletes, and significantly decreased sports anxiety and experiential avoidance in athletes. Lastly, Vveinhardt and Kaspare (2022) analyzed the relationship between mindfulness practices and the psychological state and qualification of 371 kyokushin karate athletes. Via a survey including the Mindful Attention Awareness Scale (MAAS-15) and the Depression, Anxiety, and Stress Scale (DASS-21), they found a positive influence of mindfulness on psychological state and enhanced athletic performance, as well as a moderate negative correlation showing a reduction of stress and anxiety via mindfulness. They concluded that mindfulness practices are an effective tool to improve the physical and psychological state of athletes in preparation for competitions, by reducing depression, anxiety, and stress levels, and improving attention (Vveinhardt and Kaspare, 2022). Considering these effects and their potential impact on athletes' well-being as well as performance increase by boosting attentional control, the findings of this thesis suggest that only one brief session before an important event is probably insufficient to elicit the desired effects.

7.4.3 Research

There are several possible implications for future research:

First, the findings of this thesis should be replicated and confirmed, especially regarding some of the limitations of our experiments. As these studies were one of the first to investigate the described effects in chronometric mental rotation, similar experiments could add weight to support or also contradict our findings. The possibility of creating a positive impact on cognitive performance due to a brief one-time intervention could have many beneficial implications. As described above, useful practices and routines could be developed to teach to pupils and students to exercise prior to exams and tests. It could also be beneficial in sports. Therefore, future research could create, vary, and compare different practices and techniques to establish an impactful one-time intervention to induce the wanted effects.

Second, coming studies could also focus on identifying the amount of necessary mindfulness sessions to elicit effective changes consistently and significantly in behavioral performance and cognitive load. For instance, using a six-week yoga and mindfulness intervention program, Lemay et al. (2019) found reduced stress and anxiety levels in college students and concluded that adopting such a mindfulness practice only once a week may have a positive impact on stress and anxiety reduction. Accordingly, future mental rotation studies could involve groups with one to four weekly meditation sessions over two to six weeks. At the start and end of a fixed time interval, all groups would complete a chronometric mental rotation test with a measure for cognitive load in a pre-post-test design. One general accomplishment of this endeavor could be an adequately short program for non-meditators to participate in to induce effective changes in relatively little time. They could benefit from the positive effects short-term and might also consider incorporating mindfulness more long-term into their lifestyle.

Third, newly developed programs and interventions should be validated and compared to other forms of practices. For instance, MacCoon et al. (2012) compared an active control condition (i.e., Health Enhancement Program) to an MBSR program in a clinical context. Their purpose was to isolate and test mindfulness as an active ingredient, opposed to other studies that solely rely on wait-list control comparisons. Accordingly, different forms of embodied mindfulness as well as other possibilities should be further investigated and compared in the future. For example, Repo et al. (2022) compared two mindfulness program groups and a control group over the course of an academic year and found beneficial effects for—but no difference between—the two programs and did find that practicing mindfulness at least twice a week seems beneficial to reduce psychological stress in students long-term. Like the mentioned studies and approaches, future research on the effects of embodied cognition on chronometric mental rotation performance could conduct studies to verify the efficacy of different mindfulness programs and include variation regarding the programmed time schedules.

Fourth, the analysis of sex differences requires further investigation, which also includes the similarities and differences with gender differences, and whether they should be regarded as the same. As there are many reasons for possible sex differences (e.g., biological, sociocultural, and cognitive ones), they should be further analyzed and compared for different sample groups, including a variety of genders (e.g., cisgender and transgender; see Burgund, 2021). For chronometric tests, this may also include variations in the test design like different time limits or variations of stimuli.

Fifth, the possible reduction of sex differences via training should be further investigated—also in combination with the effects that are in focus of this thesis. To date, only the study of Geng et al. (2011) exists, where it was shown that participants of both sexes responded faster after mindful than mindless learning. These effects as well as the general training effects in mental rotation should be scrutinized in the future. Next to the observed improvements in speed and accuracy rates in the mental rotation tasks, it would be interesting to analyze cognitive load over multiple learning sessions. For example, participants might be able to focus better and thus dedicate more cognitive effort into a single task, but also, overall cognitive load could be reduced, as our results of the second and third study indicate.

Sixth, expanding on the findings of the first study, future studies could further analyze variations of the trial layout (e.g., Jost & Jansen, 2020) or the stimulus material (e.g., Amorim et al., 2006). This could incorporate different figures (e.g., with varying difficulty) or figure versions (e.g., with multiple levels of embodiment), as well as different coloring patterns (e.g., Rahe et al., 2022). The latter was tested for stereotype activation by stimulus color in psychometric mental rotation tests and revealed non-significant effects on the performance (Rahe et al., 2022). To build on and expand this knowledge, these findings could be replicated, verified, and adapted for chronometric mental rotation performance, while also including sex differences and cognitive load analysis. An interesting follow-up study to this thesis could involve the analysis of how the embodiment effects of mental rotation stimulus material and mindfulness intervention might interact with each other. For example, the effects might be augmented in combination, or a ceiling effect might be observed, which could indicate a limited facilitation effect of embodiment in this cognitive task. Lastly, this analysis could include sex differences. Another follow-up study to this thesis could be to test the embodiment effects of (various) embodied stimuli in interaction with stereotype threat induction. This approach could also be followed by integrating different forms of embodiment effects (e.g., mindfulness), sex differences, and cognitive load measurements. Additionally, new analyses could include pupillometry, but also other measures of neurological activity like functional near-infrared spectroscopy (fNIRS) and electroencephalography (EEG).

Seventh, in addition to the impact of treatment design, stimuli differences, learning effects, and gender differences, research should verify whether the improvements relate and transfer to other spatial abilities and cognitive tasks connected to STEM performance, with respect to the aforementioned gap in gender equality in the STEM fields.

7.5 Conclusion

In the three presented studies, we investigated the effects of embodiment—via embodied stimuli and embodied mindfulness—and stereotype threat on chronometric mental rotation task performance and cognitive load. We also analyzed possible sex differences and provided evidence that the effects elicited by embodied test stimuli as well as one-time mindfulness interventions affects men and women similarly in their behavioral performance and cognitive load. Regarding these variables, the thesis's findings also indicate that one brief mindfulness intervention may not be enough to induce significant embodiment effects. With the addition of a stereotype threat induction, the results of the third study indicate that, generally, higher levels of state mindfulness could prevent stereotype threat to influence cognitive performance. Lastly, the first study shows that the mental rotation of embodied figures is easier to process and needs less cognitive effort to solve the task.

The overall findings provide possible directions for future research as well as practical relevance for education and sports. For the analysis of spatial abilities and cognitive performance, future studies could vary the aspects of the tests—using figures that differ in difficulty, embodiment, or color pattern—and apply other methods of neurological measurements. Furthermore, research should investigate, refine, and validate short effective programs to facilitate the usage of embodied mindfulness for non-meditators and laymen. These effective and efficient programs could then be integrated at schools and universities to improve emotion regulation and attentional control. While being a complex system, a possible long-term benefit could be an increased reduction of the gender gap in STEM fields.

In conclusion, this thesis provides new insights about possibly beneficial aspects of embodied cognition and mindfulness on mental rotation ability and cognitive load, and addresses sex differences and the influence of stereotype threat. The results may improve and inspire future research, and the discussion may highlight the practical importance due to the connection between mental rotation, mathematical abilities, and employment in the STEM fields.

8 Declarations

8.1 Ethical Standards

All experiments were conducted according to ethical declaration of Helsinki. I communicated all considerations necessary to assess the question of ethical legitimacy of the studies.

8.2 Informed Consent to Participate and to Publish

Informed consent to participate and to publish anonymous results was obtained from all individual participants included in all studies.

8.3 Acknowledgements

I want to thank the whole research team of the Institute of Sport Science for the thorough and helpful discussions about all studies. I also thank my research participants, who are the basis for the existence of my experimental work. Furthermore, I want to thank the following members of the University of Regensburg for their specific support in some of the studies. I thank Prof. Dr. Kenneth Holmqvist (University of Lund) for his professional advice (Study 1 and 3) and Prof. Dr. Mark W. Greenlee for providing us with a spot meter (Study 1 and 3). Nils Barthelme wrote the Presentation® script (Study 1 and 3). Alexander Kalus edited the Presentation® script (Study 3) and wrote the OpenSesame script (Study 2). Data collection was supported by Bianca Günther (Study 1), Melinda Herfet (Study 1), Stefan Gruber (Study 1 and 3), Dr. Philipp Hofmann (Study 3), and Alexander Kalus (Study 3). I thank Dr. Leonardo Jost for his support with data analysis and co-authoring (Study 1 and 3). I thank Bianca Günther for her support with study conception and design, and co-authoring (Study 1).

8.4 Open Research Practices

The data that support the findings of the third study are openly available in “zenodo” at https://doi.org/10.5281/zenodo.4081753.

The data that support the findings of the first and second study, as well as the Presentation® (Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com), OpenSesame (Mathôt et al., 2012), and R (R Core Team, 2018) code used for the studies are available from the corresponding author (robertbauer13@gmail.com) upon reasonable request.

8.5 Competing Interests

There were no competing interests for any of the studies.

8.6 Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Aaron, R. V., Blain, S. D., Snodgress, M. A., & Park, S. (2020). Quadratic relationship between alexithymia and interoceptive accuracy, and results from a pilot mindfulness intervention. *Frontiers in Psychiatry, 11*, 132. https://doi.org/10.3389/fpsyt.2020.00132

Aboyoun, D. C., & Dabbs, J. M. (1998). The Hess pupil dilation findings: Sex or novelty?. *Social Behavior and Personality: an international journal, 26*(4), 415-419. https://doi.org/10.2224/sbp.1998.26.4.415

Alexander, C. N., Langer, E. J., Newman, R. I., Chandler, H. M., & Davies, J. L. (1989). Transcendental meditation, mindfulness, and longevity: an experimental study with the elderly. *Journal of Personality and Social Psychology, 57*(6), 950–964. https://doi.org/10.1037/0022-3514.57.6.950

Alexander, G. M., & Son, T. (2007). Androgens and eye movements in women and men during a test of mental rotation ability. *Hormones and Behavior, 52*(2), 197-204. https://doi.org/10.1016/j.yhbeh.2007.01.011

Amorim, M.-A., Isableu, B., & Jarraya, M. (2006). Embodied spatial transformations: "Body analogy" for the mental rotation of objects. *Journal of Experimental Psychology: General, 135*(3), 327–347. https://doi.org/10.1037/0096-3445.135.3.327

Amundsen, R., Riby, L. M., Hamilton, C., Hope, M., & McGann, D. (2020). Mindfulness in primary school children as a route to enhanced life satisfaction, positive outlook and effective emotion regulation. *BMC Psychology, 8*(1), 71. https://doi.org/10.1186/s40359-020-00428-y

Andreassi, J. L. (2000). Pupillary response and behavior. In J. L. Andreassi (Ed.), *Psychophysiology: Human behavior & physiological response* (pp. 218–233): Mahwah, NJ: Lawrence Erlbaum Associates.

Anguera, J. A., Reuter-Lorenz, P. A., Willingham, D. T., & Seidler, R. D. (2010). Contributions of spatial working memory to visuomotor learning. *Journal of cognitive neuroscience, 22*(9), 1917-1930. https://doi.org/10.1162/jocn.2009.21351

Aronson, J., & McGlone, M. S. (2009). Stereotype and social identity threat. In T. D. Nelson (Ed.*), Handbook of prejudice, stereotyping, and discrimination* (pp. 153–178). Psychology Press, New York, NY: Taylor & Francis Group. https://doi.org/10.4324/9780203361993

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu. Rev. Neurosci., 28*, 403-450. https://doi.org/10.1146/annurev.neuro.28.061604.135709

Austin, J. H. (1998). *Zen and the brain: Toward an understanding of meditation and consciousness.* Cambridge, MA: MIT Press.

Avants, S. K., Beitel, M., & Margolin, A. (2005). Making the shift from 'addict self' to 'spiritual self': Results from a Stage I study of Spiritual Self-Schema (3-S) therapy for the treatment of addiction and HIV risk behavior. *Mental Health, Religion & Culture, 8*(3), 167-177. https://doi.org/10.1080/13694670500138924

Ayres, P. (2006). Impact of reducing intrinsic cognitive load on learning in a mathematical domain. Applied Cognitive Psychology: *The Official Journal of the Society for Applied Research in Memory and Cognition, 20*(3), 287-298. https://doi.org/10.1002/acp.1245

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Baddeley, A. (1992). Working memory. *Science, 255*, 556–559. https://doi.org/10.1126/science.1736359

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science, 4*, 417–423. https://doi.org/10.1016/S1364-6613(00)01538-2

Baenninger, M., & Newcombe, N. (1989). The role of experience in spatial test performance: A meta-analysis. *Sex roles, 20*(5-6), 327-344. https://doi.org/10.1007/BF00287729

Baer, R. A., Smith, G. T., & Allen, K. B. (2004). Assessment of mindfulness by self-report: The Kentucky Inventory of Mindfulness Skills. *Assessment, 11*(3), 191-206. https://doi.org/10.1177/1073191104268029

Baguley, T. (2009). Standardized or simple effect size: What should be reported? *British Journal of Psychology, 100*(3), 603–617. https://doi.org/10.1348/000712608X377117

Bamber, M. D., & Schneider, J. K. (2016). Mindfulness-based meditation to decrease stress and anxiety in college students: A narrative synthesis of the research. *Educational Research Review, 18*, 1-32. https://doi.org/10.1016/j.edurev.2015.12.004

Barkus, E. (2020). Effects of working memory training on emotion regulation: Transdiagnostic review. *PsyCh journal, 9*(2), 258-279. https://doi.org/10.1002/pchj.353

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*(1), 617–645. https://doi.org/10.1146/annurev.psych.59.103006.093639

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). *Parsimonious Mixed Models.* ArXiv:1506.04967. http://arxiv.org/abs/1506.04967

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48. https://doi.org/10.18637/jss.v067.i01

Bauer, R., Jost, L., & Jansen, P. (2021). The effect of mindfulness and stereotype threat in mental rotation: a pupillometry study. *Journal of Cognitive Psychology, 33*(8), 861-876. https://doi.org/10.1080/20445911.2021.1967366

Bauer, R., Jost, L., Günther, B., & Jansen, P. (2022). Pupillometry as a measure of cognitive load in mental rotation tasks with abstract and embodied figures. *Psychological Research, 86*(5), 1382-1396. https://doi.org/10.1007/s00426-021-01568-5

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (p. 142–162). Cambridge University Press.

Beilock, S. L., Rydell, R. J., & McConnell, A. R. (2007). Stereotype threat and working memory: Mechanisms, alleviation, and spillover. *Journal of Experimental Psychology: General, 136*(2), 256–276. https://doi.org/10.1037/0096-3445.136.2.256

Bhatia, N., & Bhatia, S. (2021). Changes in Gender Stereotypes Over Time: A Computational Analysis. *Psychology of Women Quarterly, 45*(1), 106–125. https://doi.org/10.1177/0361684320977178

Binda, P., Pereverzeva, M., & Murray, S. O. (2013). Attention to bright surfaces enhances the pupillary light reflex. *Journal of Neuroscience, 33*(5), 2199–2204. https://doi.org/10.1523/JNEUROSCI.3440-12.2013

Bishop, P. A., & Herron, R. L. (2015). Use and Misuse of the Likert Item Responses and Other Ordinal Measures. *International Journal of Exercise Science, 8*(3), 297–302. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4833473/

Blascovich, J., Spencer, S. J., Quinn, D., & Steele, C. (2001). African Americans and high blood pressure: The role of stereotype threat. *Psychological science, 12*(3), 225-229. https://doi.org/10.1111/1467-9280.00

Bochynska, A., Postma, A., Vulchanova, M., & Laeng, B. (2021). More mental rotation time does not imply more mental effort: Pupillary diameters do not change with angular distance. *Brain and cognition, 148*, 105670. https://doi.org/10.1016/j.bandc.2020.105670

Bokk, O., & Forster, B. (2022). The effect of a short mindfulness meditation on somatosensory attention. *Mindfulness, 13*(8), 2022-2030. https://doi.org/10.1007/s12671-022-01938-z

Booth, J. R., MacWhinney, B., Thulborn, K. R., Sacco, K., Voyvodic, J. T., & Feldman, H. M. (2000). Developmental and lesion effects in brain activation during sentence comprehension and mental rotation. *Developmental neuropsychology, 18*(2), 139-169. https://doi.org/10.1207/S15326942DN1802_1

Bowen, S., Witkiewitz, K., Dillworth, T. M., Chawla, N., Simpson, T. L., Ostafin, B. D., Larimer, M. E., Blume, A. W., Parks, G. A., & Marlatt, G. A. (2006). Mindfulness meditation and substance use in an incarcerated population. *Psychology of Addictive Behaviors, 20*(3), 343–347. https://doi.org/10.1037/0893-164X.20.3.343

Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods, 23*(3), 389–411. https://doi.org/10.1037/met0000159

Breda, T., Jouini, E., Napp, C., & Thebault, G. (2020). Gender stereotypes can explain the gender-equality paradox. *Proceedings of the National Academy of Sciences, 117*(49), 31063-31069. https://doi.org/10.1073/pnas.2008704117

Brehm, J. W., & Self, E. A. (1989). The intensity of motivation. *Annual review of psychology, 40*(1), 109-131. https://doi.org/10.1146/annurev.ps.40.020189.000545

Brown, K. W., & Ryan, R. M. (2003). The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology, 84*(4), 822–848. https://doi.org/10.1037/0022-3514.84.4.822

Brysbaert, M., & Stevens, M. (2018). Power Analysis and Effect Size in Mixed Effects Models: A Tutorial. *Journal of Cognition, 1*(1), 9. https://doi.org/10.5334/joc.10

Bryson, B. (2004). *Eine kurze Geschichte von fast allem [a short history of nearly everything]*. Goldmann Verlag.

Buccino, G., Lui, F., Canessa, N., Patteri, I., Lagravinese, G., Benuzzi, F., … & Rizzolatti, G. (2004). Neural circuits involved in the recognition of actions performed by non-conspecifics: an fMRI study. *Journal of Cognitive Neuroscience, 16*(1), 114–126. https://doi.org/10.1162/089892904322755601

Buckley, J., Seery, N., & Canty, D. (2018). A Heuristic Framework of Spatial Ability: a Review and Synthesis of Spatial Factor Literature to Support its Translation into STEM Education. *Educational Psychology Review, 30*(3), 947–972. https://doi.org/10.1007/s10648-018-9432-z

Burgund, E. D. (2021). Task-domain and hemisphere-asymmetry effects in cisgender and transmale individuals. *Plos one, 16*(12), e0260542. https://doi.org/10.1371/journal.pone.0260542

Büsch, D., Pabst, J., Naundorf, F., Braun, J., Marschall, F., Schumacher, K., Wilhelm, A., & Granacher, U. (2015). Subjektive Beanspruchung im Krafttraining. In U. Granacher (Ed.), *Krafttraining: "Kraftvoll durchs Leben“: Jahrestagung der dvs-Sektion Trainingswissenschaft vom 28.-30. Mai 2015 in Potsdam (Abstractband)* (p. 13). Potsdam: Uni-Print.

Cahn, B. R., & Polich, J. (2006). Meditation states and traits: EEG, ERP, and neuroimaging studies. *Psychological Bulletin, 132*(2), 180–211. https://doi.org/10.1037/0033-2909.132.2.180

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: an FMRI study with expert dancers. *Cerebral Cortex, 15*(8), 1243–1249. https://doi.org/10.1093/cercor/bhi007

Campaña, J. C., Gimenez-Nadal, J. I., & Velilla, J. (2023). Measuring Gender Gaps in Time Allocation in Europe. *Social Indicators Research, 165*(2), 519-553. https://doi.org/10.1007/s11205-022-03026-0

Campbell, M. J., Toth, A. J., & Brady, N. (2018). Illuminating sex differences in mental rotation using pupillometry. *Biological Psychology, 138*, 19–26. https://doi.org/10.1016/j.biopsycho.2018.08.003

Carpenter, P. A., Just, M. A., Keller, T. A., Eddy, W., & Thulborn, K. (1999). Graded functional activation in the visuospatial system with the amount of task demand. *Journal of Cognitive Neuroscience, 11*(1), 9–24. https://doi.org/10.1162/089892999563210

Cauller, L. (1995). Layer I of primary sensory neocortex: where top-down converges upon bottom-up. *Behavioural Brain Research, 71*(1–2), 163–170. https://doi.org/10.1016/0166-4328(95)00032-1

Chandler, J., & Schwarz, N. (2009). How extending your middle finger affects your perception of others: learned movements influence concept accessibility. *Journal of Experimental Social Psychology, 45*(1), 123–128. https://doi.org/10.1016/j.jesp.2008.06.012

Chiesa, A., Calati, R., & Serretti, A. (2011). Does mindfulness training improve cognitive abilities? A systematic review of neuropsychological findings. *Clinical psychology review, 31*(3), 449-464. https://doi.org/10.1016/j.cpr.2010.11.003

Chiesa, A., Serretti, A., & Jakobsen, J. C. (2013). Mindfulness: top-down or bottom-up emotion regulation strategy? *Clinical Psychology Review, 33*(1), 82–96. https://doi.org/10.1016/j.cpr.2012.10.006

Christensen, R. H. B. (2015). ordinal—regression models for ordinal data. *R package version, 28*, 2015.

Clewett, D., Gasser, C., & Davachi, L. (2020). Pupil-linked arousal signals track the temporal organization of events in memory. *Nature Communications, 11*(1), 4007. https://doi.org/10.1038/s41467-020-17851-9

Coelho, H. F., Canter, P. H., & Ernst, E. (2013). Mindfulness-based cognitive therapy: Evaluating current evidence and informing future research. *Psychology of Consciousness: Theory, Research, and Practice, 1*(S), 97–107. https://doi.org/10.1037/2326-5523.1.S.97

Cohen, M. S., Kosslyn, S. M., Breiter, H. C., DiGirolamo, G. J., Thompson, W. L., Anderson, A. K., Bookheimer, S. Y., Rosen, B. R., & Belliveau, J. W. (1996). Changes in cortical activity during mental rotation A mapping study using functional MRI. *Brain, 119(1)*, 89–100. https://doi.org/10.1093/brain/119.1.89

Colzato, L. S., van der Wel, P., Sellaro, R., & Hommel, B. (2016). A single bout of meditation biases cognitive control but not attentional focusing: Evidence from the global–local task. *Consciousness and Cognition, 39*, 1-7. https://doi.org/10.1016/j.concog.2015.11.003

Coyne, J. T., Foroughi, C., & Sibley, C. (2017). Pupil Diameter and Performance in a Supervisory Control Task: A Measure of Effort or Individual Differences? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 61*(1), 865–869. https://doi.org/10.1177/1541931213601689

Croizet, J. C., Després, G., Gauzins, M. E., Huguet, P., Leyens, J. P., & Méot, A. (2004). Stereotype threat undermines intellectual performance by triggering a disruptive mental load. *Personality and social psychology bulletin, 30*(6), 721-731. https://doi.org/10.1177/0146167204263961

Datta, A., Cusack, R., Hawkins, K., Heutink, J., Rorden, C., Robertson, I. H., & Manly, T. (2007). The P300 as a marker of waning attention and error propensity. *Computational Intelligence and Neuroscience, 2007*, Article ID 093968. https://doi.org/10.1155/2007/93968

Davis, K. M., Lau, M. A., & Cairns, D. R. (2009). Development and preliminary validation of a trait version of the Toronto Mindfulness Scale. *Journal of Cognitive Psychotherapy, 23*(3), 185-197. https://doi.org/10.1891/0889-8391.23.3.185

Decety, J. (2002). Is there such a thing as functional equivalence between imagined, observed, and executed action? In A. N. Meltzoff & W. Prinz (Eds.), *The imitative mind: Development, evolution, and brain bases* (p. 291–310). Cambridge University Press. https://doi.org/10.1017/CBO9780511489969.017

Decker, A., Finn, A., & Duncan, K. (2020). Errors lead to transient impairments in memory formation. *Cognition, 204*, Article 104338. https://doi.org/10.1016/j.cognition.2020.104338

Dehghani, M., Saf, A. D., Vosoughi, A., Tebbenouri, G., & Zarnagh, H. G. (2018). Effectiveness of the mindfulness-acceptance-commitment-based approach on athletic performance and sports competition anxiety: A randomized clinical trial. *Electronic physician, 10*(5), 6749. https://doi.org/10.19082/6749

Dijkstra, K., & Post, L. (2015). Mechanisms of embodiment. *Frontiers in psychology, 6*, 1525. https://doi.org/10.3389/fpsyg.2015.01525

Doyle, R. A., & Voyer, D. (2016). Stereotype manipulation effects on math and spatial test performance: A meta-analysis. *Learning and Individual Differences, 47*, 103-116. https://doi.org/10.1016/j.lindif.2015.12.018

Ducrocq, E., Wilson, M., Vine, S., & Derakshan, N. (2016). Training attentional control improves cognitive and motor task performance. *Journal of sport and exercise psychology, 38*(5), 521-533. https://doi.org/10.1123/jsep.2016-0052

Ekman, P., Davidson, R. J., Ricard, M., & Alan Wallace, B. (2005). Buddhist and psychological perspectives on emotions and well-being. *Current Directions in Psychological Science, 14*(2), 59-63. https://doi.org/10.1111/j.0963-7214.2005.003

Eldar, E., Cohen, J. D., & Niv, Y. (2013). The effects of neural gain on attention and learning. *Nature Neuroscience, 16*(8), 1146–1153. https://doi.org/10.1038/nn.3428

Epel, E., Daubenmier, J., Moskowitz, J. T., Folkman, S., & Blackburn, E. (2009). Can meditation slow rate of cellular aging? Cognitive stress, mindfulness, and telomeres. *Annals of the new York Academy of Sciences, 1172*(1), 34-53. https://doi.org/10.1111/j.1749-6632.2009.04414.x

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

Feingold, A. (2009). Effect sizes for growth-modeling analysis for controlled clinical trials in the same metric as for classical analysis. *Psychological Methods, 14*(1), 43–53. https://doi.org/10.1037/a0014699

Fischer, R. (1971). A Cartography of the Ecstatic and Meditative States: The experimental and experiential features of a perception-hallucination continuum are considered. *Science, 174*(4012), 897-904. https://doi.org/10.1126/science.174.4012.8

Ford, T. E., Ferguson, M. A., Brooks, J. L., & Hagadone, K. M. (2004). Coping sense of humor reduces effects of stereotype threat on women's math performance. *Personality and Social Psychology Bulletin, 30*(5), 643-653. https://doi.org/10.1177/0146167203262851

French, J., Ekstrom, R., & Price, L. (1963). Kit of reference tests for cognitive factors. Princeton, NJ: Educational Testing Service.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The Structure of Working Memory From 4 to 15 Years of Age. *Developmental Psychology, 40*(2), 177–190. https://doi.org/10.1037/0012-1649.40.2.177

Gauthier, I., Hayward, W. G., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (2002). BOLD activity during mental rotation and viewpoint-dependent object recognition. *Neuron, 34*(1), 161-171. https://doi.org/10.1016/S0896-6273(02)00622-0

Geary, D. C., Saults, S. J., Liu, F., & Hoard, M. K. (2000). Sex Differences in Spatial Cognition, Computational Fluency, and Arithmetical Reasoning. *Journal of Experimental Child Psychology, 77*(4), 337–353. https://doi.org/10.1006/jecp.2000.2594.

Geng, L., Zhang, L., & Zhang, D. (2011). Improving spatial abilities through mindfulness: effects on the mental rotation task. *Consciousness and cognition, 20*(3), 801-806. https://doi.org/10.1016/j.concog.2011.02.004

Gilbert, J. K. (2005). *Visualization: A Metacognitive Skill in Science and Science Education*. In: Gilbert, J. K. (eds) Visualization in Science Education. Models and Modeling in Science Education, vol 1. Springer, Dordrecht. https://doi.org/10.1007/1-4020-3613-2_2

Gilzenrat, M. S., Nieuwenhuis, S., Jepma, M., & Cohen, J. D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, Affective, & Behavioral Neuroscience, 10*(2), 252–269. https://doi.org/10.3758/CABN.10.2.252

Goodmon, L. B., Brown, K., Edwards, L., Hurley, K., Hartzell, K., & Powell, T (2019). Ready, Set, Rotate: The Relationship Between Working Memory Capacity and Mental Rotation Speed. *Journal of Scientific Psychology*. Retrieved from: http://psyencelab.com/uploads/5/4/6/5/54658091/ms101pdf.pdf

Gotink, R. A., Chu, P., Busschbach, J. J. V., Benson, H., Fricchione, G. L., & Hunink, M. G. M. (2015). Standardised mindfulness-based interventions in healthcare: an overview of systematic reviews and meta-analyses of RCTs. *PloS One, 10*(4), e0124344. https://doi.org/10.1371/journal.pone.0124344

Granholm, E., Asarnow, R. F., Sarkin, A. J., & Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology, 33*(4), 457-461. https://doi.org/10.1111/j.1469-8986.1996.tb01071.x

Grant, A. M., Langer, E. J., Falk, E., & Capodilupo, C. (2004). Mindful creativity: drawing to draw distinctions. *Creativity Research Journal, 16*(2–3), 261–265. https://doi.org/10.1080/10400419.2004.9651457

Grimshaw, G. M., Sitarenios, G., & Finegan, J. A. K. (1995). Mental rotation at 7 years-relations with prenatal testosterone levels and spatial play experiences. *Brain and cognition, 29*(1), 85-100. https://doi.org/10.1006/brcg.1995.1269

Grossman, P., Niemann, L., Schmidt, S., & Walach, H. (2004). Mindfulness-based stress reduction and health benefits: a meta-analysis. *Journal of Psychosomatic Research, 57*(1), 35–43. https://doi.org/10.1016/S0022-3999%2803%2900573-7

Groves, N. B., Kofler, M. J., Wells, E. L., Day, T. N., & Chan, E. S. (2020). An examination of relations among working memory, ADHD symptoms, and emotion regulation. *Journal of abnormal child psychology, 48*(4), 525-537. https://doi.org/10.1007/s10802-019-00612-8

Halpern, D. F. (1989). The disappearance of cognitive gender differences: What you see depends on where you look. *American Psychologist, 44*(8), 1156–1158. https://doi.org/10.1037/0003-066X.44.8.1156

Halpern, D. F. (2013). *Sex differences in cognitive abilities* (4th ed.). Psychology Press, Taylor & Francis.

Hanel, P. H., & Vione, K. C. (2016). Do student samples provide an accurate estimate of the general public?. *PloS one, 11*(12), e0168354. https://doi.org/10.1371/journal.pone.0168354

Harle, M., & Towns, M. (2011). A review of spatial ability literature, its connection to chemistry, and implications for instruction. *Journal of Chemical Education, 88*(3), 351–360. https://doi.org/10.1021/ed900003n

Hart, R., Ivtzan, I., & Hart, D. (2013). Mind the gap in mindfulness research: a comparative account of the leading schools of thought. *Review of General Psychology, 17*(4), 453–466. https://doi.org/10.1037/a0035212

Hassed, C. (2021). The art of introducing mindfulness into medical and allied health curricula. *Mindfulness, 12*(8), 1909-1919. https://doi.org/10.1007/s12671-021-01647-z

Hausmann, M. (2014). Arts versus science - Academic background implicitly activates gender stereotypes on cognitive abilities with threat raising men's (but lowering women's) performance. *Intelligence, 46*(1), 235–245. https://doi.org/10.1016/j.intell.2014.07.004

Hausmann, M., Schoofs, D., Rosenthal, H. E. S., & Jordan, K. (2009). Interactive effects of sex hormones and gender stereotypes on cognitive sex differences-A psychobiosocial approach. *Psychoneuroendocrinology, 34*(3), 389–401. https://doi.org/10.1016/j.psyneuen.2008.09.019

Hausmann, M., Slabbekoorn, D., Van Goozen, S. H. M., Cohen-Kettenis, P. T., & Güntürkün, O. (2000). Sex hormones affect spatial abilities during the menstrual cycle. *Behavioral Neuroscience, 114*(6), 1245–1250. https://doi.org/10.1037/0735-7044.114.6.1245

Hayes, T. R., & Petrov, A. A. (2016). Mapping and correcting the influence of gaze position on pupil size measurements. *Behavior Research Methods, 48*(2), 510-527. https://doi.org/10.3758/s13428-015-0588-x

Hedges, L. V. (2007). Effect sizes in cluster-randomized designs. *Journal of Educational and Behavioral Statistics, 32*(4), 341–370. https://doi.org/10.3102/1076998606298043

Hegarty, M. (2018). Ability and sex differences in spatial thinking: What does the mental rotation test really measure? *Psychonomic Bulletin and Review, 25*(3), 1212–1219. https://doi.org/10.3758/s13423-017-1347-z

Hegarty, M., & Kozhevnikov, M. (1999). Types of visual–spatial representations and mathematical problem solving. *Journal of Educational Psychology, 91*(4), 684–689. https://doi.org/10.1037/0022-0663.91.4.684

Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence, 32*(2), 175–191. https://doi.org/10.1016/j.intell.2003.12.001

Hegarty, M., & Waller, D. A. (2005). Individual differences in spatial abilities. *The Cambridge handbook of visuospatial thinking*, 121-169. Cambridge University Press. https://doi.org/10.1017/CBO9780511610448.005

Heil, M., & Jansen-Osmann, P. (2008). Sex differences in mental rotation with polygons of different complexity: Do men utilize holistic processes whereas women prefer piecemeal ones? *Quarterly Journal of Experimental Psychology, 61*(5), 683–689. https://doi.org/10.1080/17470210701822967

Heil, M., Jansen, P., Quaiser-Pohl, C., & Neuburger, S. (2012). Gender-specific effects of artificially induced gender beliefs in mental rotation. *Learning and Individual Differences, 22*(3), 350–353. https://doi.org/10.1016/j.lindif.2012.01.004

Heil, M., & Rolke, B. (2002). Toward a chronopsychophysiology of mental rotation. *Psychophysiology, 39*(4), 414–422. https://doi.org/10.1111/1469-8986.3940414

Heppner, W. L., Kernis, M. H., Lakey, C. E., Campbell, W. K., Goldman, B. M., Davis, P. J., et al. (2008). Mindfulness as a means of reducing aggressive behaviour: Dispositional and situational evidence. *Aggressive Behavior, 34*, 486–496. https://doi.org/10.1002/ab.20258

Hershman, R., & Henik, A. (2019). Dissociation between reaction time and pupil dilation in the Stroop task. *Journal of Experimental Psychology: Learning Memory and Cognition, 45*(10), 1899-1909. https://doi.org/10.1037/xlm0000690

Hertzog, C., Vernon, M. C., & Rypma, B. (1993). Age differences in mental rotation task performance: The influence of speed/accuracy tradeoffs. *Journal of Gerontology, 48*(3), 150-156. https://doi.org/10.1093/geronj/48.3.P150

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science, 143*(3611), 1190-1192. https://doi.org/10.1126/science.143.3611.1190

Hilbert, S., Stadler, M., Lindl, A., Naumann, F., & Bühner, M. (2019). Analyzing longitudinal intervention studies with linear mixed models. *TPM - Testing, Psychometrics, Methodology in Applied Psychology, 26*(1), 101–119. https://doi.org/10.4473/TPM26.1.6

Hockey, G. R. J. (1997). Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. *Biological psychology, 45*(1-3), 73-93. https://doi.org/10.1016/S0301-0511(96)05223-4

Hofmann, S. G., Sawyer, A. T., Fang, A., & Asnaani, A. (2012). Emotion dysregulation model of mood and anxiety disorders. *Depression and Anxiety, 29*(5), 409–416. https://doi.org/10.1002/da.21888

Hofmann, S. G., Sawyer, A. T., Witt, A. A., & Oh, D. (2010). The effect of mindfulness-based therapy on anxiety and depression: *A meta-analytic review. Journal of consulting and clinical psychology, 78*(2), 169. https://doi.org/10.1037/a0018555

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford.

Hölzel, B. K., Lazar, S. W., Gard, T., Schuman-Olivier, Z., Vago, D. R., & Ott, U. (2011). How does mindfulness meditation work? Proposing mechanisms of action from a conceptual and neural perspective. *Perspectives on Psychological Science, 6*(6), 537–559. https://doi.org/10.1177/1745691611419671

Humphreys, L. G., Lubinski, D., & Yao, G. (1993). Utility of predicting group membership and the role of spatial visualization in becoming an engineer, physical scientist, or artist. *Journal of Applied Psychology, 78*(2), 250–261. https://doi.org/10.1037/0021-9010.78.2.250

Hustá, C., Dalmaijer, E., Belopolsky, A., & Mathôt, S. (2019). The pupillary light response reflects visual working memory content. *Journal of Experimental Psychology: Human Perception and Performance, 45*(11), 1522–1528. https://doi.org/10.1037/xhp0000689

Hyun, J. S., & Luck, S. J. (2007). Visual working memory as the substrate for mental rotation. *Psychonomic Bulletin and Review, 14*(1), 154–158. https://doi.org/10.3758/BF03194043

Inzlicht, M., Aronson, J., Good, C., & McKay, L. (2006). A particular resiliency to threatening environments. *Journal of experimental social psychology, 42*(3), 323-336. https://doi.org/10.1016/j.jesp.2005.05.005

Inzlicht, M., McKay, L., & Aronson, J. (2006). Stigma as ego depletion: How being the target of prejudice affects self-control. *Psychological Science, 17*(3), 262-269. https://doi.org/10.1111/j.1467-9280.2006.01695.x

Isreal, J. B., Chesney, G. L., Wickens, C. D., & Donchin, E. (1980). P300 and tracking difficulty: Evidence for multiple resources in dual-task performance. *Psychophysiology, 17*(3), 259-273. https://doi.org/10.1111/j.1469-8986.1980.tb00146.x

Jackson, I., & Sirois, S. (2009). Infant cognition: Going full factorial with pupil dilation. *Developmental Science, 12*, 670–679. https://doi.org/10.1111/j.1467-7687.2008.00805.x

Jansen-Osmann, P., & Heil, M. (2007). Suitable stimuli to obtain (no) gender differences in the speed of cognitive processes involved in mental rotation. *Brain and Cognition, 64*(3), 217–227. https://doi.org/10.1016/j.bandc.2007.03.002

Jansen, P., Lehmann, J., & Van Doren, J. (2012). Mental rotation performance in male soccer players. *PloS one, 7*(10), e48620. https://doi.org/10.1371/journal.pone.0048620

Jansen, P., Render, A., Scheer, C., & Siebertz, M. (2020). Mental rotation with abstract and embodied objects as stimuli: Evidence from event-related potential (ERP). *Experimental Brain Research, 238*, 525-535. https://doi.org/10.1007/s00221-020-05734-w

Jepma, M., & Nieuwenhuis, S. (2011). Pupil diameter predicts changes in the exploration–exploitation trade-off: Evidence for the adaptive gain theory. *Journal of cognitive neuroscience, 23*(7), 1587-1596. https://doi.org/10.1162/jocn.2010.21548

Johns, M., Inzlicht, M., & Schmader, T. (2008). Stereotype threat and executive resource depletion: Examining the influence of emotion regulation. *Journal of Experimental Psychology: General, 137*(4), 691–705. https://doi.org/10.1037/a0013834

Johnson, S. P., & Moore, D. S. (2020). Spatial thinking in infancy: Origins and development of mental rotation between 3 and 10 months of age. *Cognitive research: principles and implications, 5*, 1-14. https://doi.org/10.1186/s41235-020-00212-x

Jolicœur, P., Regehr, S., Smith, L. B. J. P., & Smith, G. N. (1985). Mental rotation of representations of two-dimensional and three-dimensional objects. *Canadian Journal of Psychology/Revue Canadienne de Psychologie, 39*(1), 100–129. https://doi.org/10.1037/h0080118

Jordan, K., Heinze, H. J., Lutz, K., Kanowski, M., and Jäncke, L. (2001). Cortical activations during the mental rotation of different visual objects. *Neuroimage 13*, 143–152. https://doi.org/10.1006/nimg.2000.0677

Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. *Neuron, 89*(1), 221–234. https://doi.org/10.1016/j.neuron.2015.11.028

Jost, L. (2022). *Mental Rotation-The Test Design, Sex Differences, and the Link to Physical Activity* (Doctoral dissertation). https://doi.org/10.5283/epub.51432

Jost, L., & Jansen, P. (2020). A novel approach to analyzing all trials in chronometric mental rotation and description of a flexible extended library of stimuli. *Spatial Cognition & Computation, 20*(3), 234–256. https://doi.org/10.1080/13875868.2020.1754833

Jost, L., Weishäupl, A., & Jansen, P. (2023). Interactions between simultaneous aerobic exercise and mental rotation. *Current Psychology, 42*(6), 4682-4695. https://doi.org/10.1007/s12144-021-01785-6

Jostmann, N. B., Lakens, D., & Schubert, T. W. (2009). Weight as an embodiment of importance. *Psychological Science, 20*(9), 1169–1174. https://doi.org/10.1111/j.1467-9280.2009.02426.x

Jürgens, R., Fischer, J., & Schacht, A. (2018). Hot speech and exploding bombs: Autonomic arousal during emotion classification of prosodic utterances and affective sounds. *Frontiers in psychology, 9*, 228. https://doi.org/10.3389/fpsyg.2018.00228

Kabat-Zinn, J. (1982). An outpatient program in behavioral medicine for chronic pain patients based on the practice of mindfulness meditation: theoretical considerations and preliminary results. *General Hospital Psychiatry, 4*(1), 33–47. https://doi.org/10.1016/0163-8343(82)90026-3

Kabat-Zinn, J. (1991). *Full catastrophe living: using the wisdom of your body and mind to face stress, pain, and illness*. NewYork, NY: Dell Publishing.

Kabat-Zinn, J. (1994). *Wherever you go, there you are: mindfulness meditation in everyday life.* New York: Hyperion.

Kabat-Zinn, J. (2003). Mindfulness-based interventions in context: past, present, and future. Clinical Psychology: *Science and Practice, 10*(2), 144–156. https://doi.org/10.1093/clipsy.bpg016

Kabat-Zinn, J., Lipworth, L., & Burney, R. (1985). The clinical use of mindfulness meditation for the self-regulation of chronic pain. *Journal of Behavioral Medicine, 8*(2), 163–190. https://doi.org/10.1007/BF00845519

Kabat-Zinn, J., Lipworth, L., Burney, R., & Sellers, W. (1986). Four-year follow-up of a meditation-based program for the self-regulation of chronic pain: treatment outcomes and compliance. *The Clinical Journal of Pain, 2*(3), 159–774.

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Englewood Cliffs, NJ: Prentice-Hall. https://doi.org/10.2307/1421603

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science, 154*(3756), 1583-1585. https://doi.org/10.1126/science.154.3756.1583

Kaufman, S. B. (2007). Sex differences in mental rotation and spatial visualization ability: Can they be accounted for by differences in working memory capacity? *Intelligence, 35*(3), 211–223. https://doi.org/10.1016/j.intell.2006.07.009

Keene, P. A., DeBettencourt, M. T., Awh, E., & Vogel, E. K. (2022). Pupillometry signatures of sustained attention and working memory. *Attention, Perception, & Psychophysics, 84*(8), 2472-2482. https://doi.org/10.3758/s13414-022-02557-5

Keller, J. (2007). Stereotype threat in classroom settings: The interactive effect of domain identification, task difficulty and stereotype threat on female students' maths performance. *British journal of educational psychology, 77*(2), 323-338. https://doi.org/10.1348/000709906X113662

Khoury, B., Knäuper, B., Pagnini, F., Trent, N., Chiesa, A., & Carrière, K. (2017). Embodied mindfulness. *Mindfulness, 8*, 1160-1171. https://doi.org/10.1007/s12671-017-0700-7

Kida, T., Kaneda, T., & Nishihira, Y. (2012). Dual-task repetition alters event-related brain potentials and task performance. Clinical Neurophysiology: *Official Journal of the International Federation of Clinical Neurophysiology, 123*(6), 1123–1130. https://doi.org/10.1016/j.clinph.2011.10.001

Kiken, L. G., Garland, E. L., Bluth, K., Palsson, O. S., & Gaylord, S. A. (2015). From a state to a trait: Trajectories of state mindfulness in meditation during intervention predict changes in trait mindfulness. *Personality and Individual differences, 81*, 41-46. https://doi.org/10.1016/j.paid.2014.12.044

King, A. P., Erickson, T. M., Giardino, N. D., Favorite, T., Rauch, S. A., Robinson, E., ... & Liberzon, I. (2013). A pilot study of group mindfulness-based cognitive therapy (MBCT) for combat veterans with posttraumatic stress disorder (PTSD). *Depression and anxiety, 30*(7), 638-645. https://doi.org/10.1002/da.22104

Koss M. C. (1986). Pupillary dilation as an index of central nervous system alpha 2-adrenoceptor activation. *Journal of pharmacological methods, 15*(1), 1–19. https://doi.org/10.1016/0160-5402(86)90002-1

Kozhevnikov, M., & Hegarty, M. (2001). A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory and Cognition, 29*(5), 745–756. https://doi.org/10.3758/BF03200477

Kozhevnikov, M., Motes, M. A., & Hegarty, M. (2007). Spatial visualization in physics problem solving. *Cognitive Science, 51*(4), 549-579. https://doi.org/10.1080/15326900701399897

Kret, M. E., & Sjak-Shie, E. E. (2019). Preprocessing pupil size data: Guidelines and code. *Behavior research methods, 51*(3), 1336-1342. https://doi.org/10.3758/s13428-018-1075-y

Kret, M. E., Tomonaga, M., & Matsuzawa, T. (2014). Chimpanzees and humans mimic pupil-size of conspecifics. *PLoS One, 9*(8), e104886. https://doi.org/10.1371/journal.pone.0104886

Kristeller, J. L., Baer, R. A., & Quillian-Wolever, R. (2006). Mindfulness-based approaches to eating disorders. In: Baer, RA, editor. *Mindfulness-based treatment approaches: Clinician's guide to evidence base and applications*, 75-91. San Diego: Elsevier.

Laeng, B., & Sulutvedt, U. (2014). The eye pupil adjusts to imaginary light. *Psychological Science, 25*(1), 188–197. https://doi.org/10.1177/0956797613503556

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to Western thought* (Vol. 4). New York, NY: Basic Books.

Lama, D., & Berzin, A. (1997). *The Gelug/Kagyu tradition of mahamudra*. New York, NY: Snow Lion.

Lammers, W. J., & Badia, P. (1989). Habituation of P300 to target stimuli. *Physiology and Behavior, 45*(3), 595–601. https://doi.org/10.1016/0031-9384(89)90079-6

Lange, K., Kühn, S., & Filevich, E. (2015). "Just another tool for online studies" (JATOS): An easy solution for setup and management of web servers supporting online studies. *PloS one, 10*(6), e0130834. https://doi.org/10.1371/journal.pone.0130834

Langer, E. J. (1989). *Mindfulness.* Cambridge, MA: Da Capo Press.

Langer, E. J. (1997). *The power of mindful learning.* Reading, MA: Addison Wesley Longman.

Langer, E. J. (2009). *Counterclockwise: mindful health and the power of possibility.* New York, NY: Ballantine Books.

Langer, E. J. (2011). A mindful alternative to the mind/body problem. *Journal of Cosmology, 14*.

Langer, E. J. (2012). The mindless use of medical data. *Journal of Business Research, 65*(11), 1651–1653. https://doi.org/10.1016/j.jbusres.2011.02.018

Langer, E. J., Blank, A., & Chanowitz, B. (1978). The mindlessness of ostensibly thoughtful action: The role of "placebic" information in interpersonal interaction. *Journal of Personality and Social Psychology, 36*(6), 635–642. https://doi.org/10.1037/0022-3514.36.6.635

Langer, E. J., Hatem, M., Joss, J., & Howell, M. (1989). Conditional teaching and mindful learning. *Creativity Research Journal, 2*(3), 139–150. https://doi.org/10.1080/10400418909534311

Langer, E. J., & Imber, L. (1979). When practice makes imperfect: the debilitating effects of overlearning. *Journal of Personality and Social Psychology, 37*(11), 2014–2025. https://doi.org/10.1037/0022-3514.37.11.2014

Langer, E. J., & Piper, A. I. (1987). The prevention of mindlessness. *Journal of Personality and Social Psychology, 53*(2), 280–287. https://doi.org/10.1037/0022-3514.53.2.280

Langer, E. J., Russel, T., & Eisenkraft, N. (2009). Orchestral performance and the footprint of mindfulness. *Psychology of Music, 37*(2), 125–136. https://doi.org/10.1177/0305735607086053

Lau, M. A., Bishop, S. R., Segal, Z. V., Buis, T., Anderson, N. D., Carlson, L., Shapiro, S., Carmody, J., Abbey, S., & Devins, G. (2006). The Toronto mindfulness scale: Development and validation. *Journal of clinical psychology, 62*(12), 1445-1467. https://doi.org/10.1002/jclp.20326

Lauer, J. E., Yhang, E., & Lourenco, S. F. (2019). The development of gender differences in spatial reasoning: A meta-analytic review. *Psychological bulletin, 145*(6), 537-565. https://doi.org/10.1037/bul0000191

Law, F., McGuire, L., Winterbottom, M., & Rutland, A. (2021). Children's gender stereotypes in STEM following a one-shot growth mindset intervention in a science museum. *Frontiers in Psychology, 12*, 641695. https://doi.org/10.3389/fpsyg.2021.641695

Lehmann, J., Quaiser-Pohl, C., & Jansen, P. (2014). Correlation of motor skill, mental rotation, and working memory in 3-to 6-year-old children. *European Journal of Developmental Psychology, 11*(5), 560-573. https://doi.org/10.1080/17405629.2014.888995

Lemay, V., Hoolahan, J., & Buchanan, A. (2019). Impact of a yoga and meditation intervention on students' stress and anxiety levels. *American journal of pharmaceutical education, 83*(5). https://doi.org/10.5688/ajpe7001

Levy, B. R., Jennings, P., & Langer, E. J. (2001). Improving attention in old age. *Journal of Adult Development, 8*(3), 189–192. https://doi.org/10.1023/A:1009546408665

Lewis, K. B., Graham, I. D., Boland, L., & Stacey, D. (2021). Writing a compelling integrated discussion: a guide for integrated discussions in article-based theses and dissertations. *International Journal of Nursing Education Scholarship, 18*(1), 20200057. https://doi.org/10.1515/ijnes-2020-0057

Linn, M. C., & Petersen, A. C. (1985). Emergence and Characterization of Sex Differences in Spatial Ability : A Meta-Analysis. *Child Development, 56*(6), 1479–1498. https://doi.org/10.2307/1130467

Liuzza, M. T., Setti, A., & Borghi, A. M. (2012). Kids observing other kids' hands: visuomotor priming in children. *Consciousness and Cognition, 21*(1), 383–392. https://doi.org/10.1016/j.concog.2011.09.015.

Loewenfeld, I. E. (1958). Mechanisms of reflex dilatation of the pupil. *Documenta Ophthalmologica, 12*(1), 185–448. https://doi.org/10.1007/BF00913471

Loewenstein, O., & Loewenfeld, I. E. (1962). The pupil. In H. Davson (Ed.), *The eye.* Vol 3. Muscular mechanisms. New York, NY: Academic Press.

Lutz, A., Dunne, J. D., & Davidson, R. J. (2007). Meditation and the neuroscience of consciousness: An introduction. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *The Cambridge handbook of consciousness* (pp. 499–551). Cambridge University Press. https://doi.org/10.1017/CBO9780511816789.020

Lutz, A., Slagter, H. A., Dunne, J. D., & Davidson, R. J. (2008). Attention regulation and monitoring in meditation. *Trends in cognitive sciences, 12*(4), 163-169. https://doi.org/10.1016/j.tics.2008.01.005

MacCoon, D. G., Imel, Z. E., Rosenkranz, M. A., Sheftel, J. G., Weng, H. Y., Sullivan, J. C., Bonus, K. A., Stoney, C. M., Salomons, T. V., Davidson, R. J., & Lutz, A. (2012). The validation of an active control intervention for Mindfulness Based Stress Reduction (MBSR). *Behaviour research and therapy, 50*(1), 3-12. https://doi.org/10.1016/j.brat.2011.10.011

Makarova, E., Aeschlimann, B., & Herzog, W. (2019). The gender gap in STEM fields: The impact of the gender stereotype of math and science on secondary students' career aspirations. In *Frontiers in Education, 4,* (p. 60). Frontiers. https://doi.org/10.3389/feduc.2019.00060

Martens, A., Johns, M., Greenberg, J., & Schimel, J. (2006). Combating stereotype threat: The effect of self-affirmation on women's intellectual performance. *Journal of Experimental Social Psychology, 42*, 236–243. https://doi.org/10.1016/j.jesp.2005.04.010

Mathôt, S. (2013). A simple way to reconstruct pupil size during eye blinks. *Available, 10*, m9. https://doi.org/10.6084/m9.figshare.688001

Mathôt, S. (2018). Pupillometry: Psychology, Physiology, and Function. *Journal of Cognition, 1*(1), 16. https://doi.org/10.5334/joc.18

Mathôt, S., Fabius, J., Van Heusden, E., & Van der Stigchel, S. (2018). Safe and sensible preprocessing and baseline correction of pupil-size data. *Behavior research methods, 50*(1), 94-106. https://doi.org/10.3758/s13428-017-1007-2

Mathôt, S., Grainger, J., & Strijkers, K. (2017). Pupillary responses to words that convey a sense of brightness or darkness. *Psychological Science, 28*(8), 1116–1124. https://doi.org/10.1177/0956797617702699

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44*(2), 314-324. https://doi.org/10.3758/s13428-011-0168-7

Mathôt, S., Van der Linden, L., Grainger, J., & Vitu, F. (2013). The pupillary light response reveals the focus of covert visual attention. *PloS one, 8*(10), e78168. https://doi.org/10.1371/journal.pone.0078168

Mathôt, S., Van der Linden, L., Grainger, J., & Vitu, F. (2015). The pupillary light response reflects eye-movement preparation. *Journal of Experimental Psychology: Human Perception and Performance, 41*(1), 28–35. https://doi.org/https://doi.org/10.1037/a0038653

Mathôt, S., & Vilotijević, A. (2022). Methods in cognitive pupillometry: Design, preprocessing, and statistical analysis. *Behavior Research Methods*, 1-23. https://doi.org/10.3758/s13428-022-01957-7

Matlin, M. W. (2011). *The psychology of women.* Cengage Learning.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language, 94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 37*(1), 1–10. https://doi.org/10.1016/j.intell.2008.08.004.

McKeering, P., & Hwang, Y. S. (2019). A systematic review of mindfulness-based school interventions with early adolescents. *Mindfulness, 10*(4), 593-610. https://doi.org/10.1007/s12671-018-0998-9

Meneghetti, C., Borella, E., & Pazzaglia, F. (2016). Mental rotation training: transfer and maintenance effects on spatial abilities. *Psychological research, 80*, 113-127. https://doi.org/10.1007/s00426-014-0644-7

Michalak, J., Burg, J., & Heidenreich, T. (2012). Don't forget your body: mindfulness, embodiment, and the treatment of depression. *Mindfulness, 3*(3), 190–199. https://doi.org/10.1007/s12671-012-0107-4

Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language, 59*(4), 475–494. https://doi.org/10.1016/j.jml.2007.11.006

Mix, K. S., Hambrick, D. Z., Satyam, V. R., Burgoyne, A. P., & Levine, S. C. (2018). The latent structure of spatial skill: A test of the 2 × 2 typology. *Cognition, 180*, 268–278. https://doi.org/10.1016/j.cognition.2018.07.012

Moè, A. (2016). Does experience with spatial school subjects favour girls' mental rotation performance? *Learning and Individual Differences, 47*, 11–16. https://doi.org/10.1016/j.lindif.2015.12.007

Moè, A. (2021). Doubling mental rotation scores in high school students: Effects of motivational and strategic trainings. *Learning and Instruction, 74*, 101461. https://doi.org/10.1016/j.learninstruc.2021.101461

Moè, A., Hausmann, M., & Hirnstein, M. (2021). Gender stereotypes and incremental beliefs in STEM and non-STEM students in three countries: Relationships with performance in cognitive tasks. *Psychological research, 85*(2), 554-567. https://doi.org/10.1007/s00426-019-01285-0

Moè, A., Jansen, P., & Pietsch, S. (2018). Childhood preference for spatial toys. Gender differences and relationships with mental rotation in STEM and non-STEM students. *Learning and Individual Differences, 68*, 108–115. https://doi.org/10.1016/j.lindif.2018.10.003

Moè, A., & Pazzaglia, F. (2006). Following the instructions!. Effects of gender beliefs in mental rotation. *Learning and Individual Differences, 16*(4), 369–377. https://doi.org/10.1016/j.lindif.2007.01.002

Moore, A., & Malinowski, P. (2009). Meditation, mindfulness and cognitive flexibility. *Consciousness and Cognition, 18*(1), 176–186. doi:10.1016/j.concog.2008.12.008

Morais Maceira, H. (2017). Economic benefits of gender equality in the EU. *Intereconomics, 52*(3), 178-183. https://doi.org/10.1007/s10272-017-0669-4

Naber, M., Alvarez, G. A., & Nakayama, K. (2013). Tracking the allocation of attention using human pupillary oscillations. *Frontiers in Psychology, 4*(919), 1–12. https://doi.org/10.3389/fpsyg.2013.00919

Nakata, H., Sakamoto, K., & Kakigi, R. (2015). Effects of task repetition on event-related potentials in somatosensory Go/No-go paradigm. *Neuroscience letters, 594*, 82-86. https://doi.org/10.1016/j.neulet.2015.03.055

Nash, J. C., & Varadhan, R. (2011). Unifying optimization algorithms to aid software system users: optimx for R. *Journal of Statistical Software, 43*(9), 1-14. https://doi.org/10.18637/jss.v043.i09

Natanzon, M., & Ferguson, M. J. (2012). Goal pursuit is grounded: the link between forward movement and achievement. *Journal of Experimental Social Psychology, 48*(1), 379–382. https://doi.org/10.1016/j.jesp.2011.06.021

Nazareth, A., Killick, R., Dick, A. S., & Pruden, S. M. (2019). Strategy selection versus flexibility: Using eye-trackers to investigate strategy use during mental rotation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45*(2), 232–245. https://doi.org/10.1037/xlm0000574

Neubauer, A. C., Bergner, S., & Schatz, M. (2010). Two-vs. three-dimensional presentation of mental rotation tasks: Sex differences and effects of training on performance and brain activation. *Intelligence, 38*(5), 529-539. https://doi.org/10.1016/j.intell.2010.06.001

Neuburger, S., Ruthsatz, V., Jansen, P., & Quaiser-Pohl, C. (2015). Can girls think spatially? Influence of implicit gender stereotype activation and rotational axis on fourth graders' mental-rotation performance. *Learning and Individual Differences, 37*, 169-175. https://doi.org/10.1016/j.lindif.2014.09.003

Newcombe, N. S., & Shipley, T. F. (2015). Thinking About Spatial Thinking: New Typology, New Assessments. In J. S. Gero (Ed.), *Studying Visual and Spatial Reasoning for Design Creativity* (pp. 179–192). Springer Netherlands. https://doi.org/10.1007/978-94-017-9297-4_10

Ngrok Software (2023). ngrok, Inc. Retrieved from: https://ngrok.com/

Nguyen, H. H. D., & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology, 93*(6), 1314–1334. https://doi.org/10.1037/a0012702

Niedenthal, P. M., Barsalou, L. W., Winkielman, P., Krauth-Gruber, S., & Ric, F. (2005). Embodiment in attitudes, social perception, and emotion. *Personality and social psychology review, 9*(3), 184-211. https://doi.org/10.1207/s15327957pspr09

Nielsen, L., & Kaszniak, A. W. (2006). Awareness of subtle emotional feelings: A comparison of long-term meditators and nonmeditators. *Emotion, 6*(3), 392–405. https://doi.org/10.1037/1528-3542.6.3.392

Nien, J. T., Wu, C. H., Yang, K. T., Cho, Y. M., Chu, C. H., Chang, Y. K., & Zhou, C. (2020). Mindfulness training enhances endurance performance and executive functions in athletes: An event-related potential study. *Neural plasticity, 2020*. Article ID 8213710. https://doi.org/10.1155/2020/8213710

Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus--norepinephrine system. *Psychological Bulletin, 131*(4), 510–532. https://doi.org/10.1037/0033-2909.131.4.510

Nieuwenhuis, S., De Geus, E. J., & Aston-Jones, G. (2011). The anatomical and functional relationship between the P3 and autonomic components of the orienting response. *Psychophysiology, 48*(2), 162–175. https://doi.org/10.1111/j.1469-8986.2010.01057.x

Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive psychology, 7*(1), 44-64. https://doi.org/10.1016/0010-0285(75)90004-3

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *Journal of Personality and Social Psychology, 83*(1), 44–59. https://doi.org/10.1037/0022-3514.83.1.44

Ostafin, B. D., & Kassman, K. T. (2012). Stepping out of history: mindfulness improves insight problem solving. *Consciousness and Cognition, 21*(2), 1031–1036. https://doi.org/10.1016/j.concog.2012.02.014

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational psychologist, 38*(1), 63-71. https://doi.org/10.1207/S15326985EP3801_8

Paas, F. G., Van Merriënboer, J. J., & Adam, J. J. (1994). Measurement of cognitive load in instructional research. *Perceptual and motor skills, 79*(1), 419-430. https://doi.org/10.2466/pms.1994.79.1.419

Pagnini, F., & Philips, D. (2015). Being mindful about mindfulness. *The Lancet Psychiatry, 2*(4), 288–289. https://doi.org/10.1016/s2215-0366(15)00041-3

Pardo-Vazquez, J. L., & Fernandez-Rey, J. (2012). Working memory capacity and mental rotation: evidence for a domain-general view. *The Spanish journal of psychology, 15*(3), 881-890. https://doi.org/10.5209/rev_SJOP.2012.v15.n3.39381

Peavler, W. S. (1974). Pupil size, information overload, and performance differences. *Psychophysiology, 11*(5), 559-566. https://doi.org/10.1111/j.1469-8986.1974.tb01114.x

Pek, J., & Flora, D. B. (2018). Reporting effect sizes in original psychological research: A discussion and tutorial. *Psychological Methods, 23*(2), 208–225. https://doi.org/10.1037/met0000126

Pennington, C. R., Litchfield, D., McLatchie, N., & Heim, D. (2019). Stereotype threat may not impact women's inhibitory control or mathematical performance: Providing support for the null hypothesis. *European Journal of Social Psychology, 49*(4), 717–734. https://doi.org/10.1002/ejsp.2540

Peters, M. (2005). Sex differences and the factor of time in solving Vandenberg and Kuse mental rotation problems. *Brain and Cognition, 57*(2), 176–184. https://doi.org/10.1016/j.bandc.2004.08.052

Peters, M., & Battista, C. (2008). Applications of mental rotation figures of the Shepard and Metzler type and description of a mental rotation stimulus library. *Brain and Cognition, 66*(3), 260–264. https://doi.org/10.1016/j.bandc.2007.09.003

Peters, M., Chisholm, P., & Laeng, B. (1995). Spatial Ability, Student Gender, and Academic Performance. *Journal of Engineering Education, 84*(1), 69–73. https://doi.org/10.1002/j.2168-9830.1995.tb00148.x

Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., & Richardson, C. (1995). A redrawn vandenberg and kuse mental rotations test - different versions and factors that affect performance. *Brain and Cognition, 28*(1), 39–58. https://doi.org/10.1006/brcg.1995.1032

Peters, M., Manning, J. T., & Reimers, S. (2007). The effects of sex, sexual orientation, and digit ratio (2D:4D) on mental rotation performance. *Archives of Sexual Behavior, 36*(2), 251–260. https://doi.org/10.1007/s10508-006-9166-8

Phillips, C. (2015). *Mindfulness in medical schools*. Retrieved from http://www.mindfulnessinstitute.ca/MindfulnessResources/MindfulnessinMedSchools.

Piaget, J., & Inhelder, B. (1969). *The psychology of the child*. Basic Books. Inc New York.

Picton, T. W. (1992). The P300 wave of the human event-related potential. *Journal of Clinical Neurophysiology, 9*(4), 456–479. https://doi.org/10.1097/00004691-199210000-00002

Pietsch, S., Jansen, P., & Lehmann, J. (2019). The choice of sports affects mental rotation performance in adolescents. *Frontiers in Neuroscience, 13*, 224. https://doi.org/10.3389/fnins.2019.00224

Polich, J. (2012). Neuropsychology of P300. In E. S. Kappenman & S. J. Luck (Eds.), *The Oxford handbook of event-related potential components*, 159-188. Oxford Handbooks Online. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.013.0089

Poock, G. K. (1973). Information processing vs pupil diameter. *Perceptual and motor skills, 37*(3), 1000-1002. https://doi.org/10.1177/003151257303700363

Portele, C., & Jansen, P. (2023). The effects of a mindfulness-based training in an elementary school in Germany. *Mindfulness, 14*(4), 830-840. https://doi.org/10.1007/s12671-023-02084-w

Quinn, P. C., & Liben, L. S. (2014). A sex difference in mental rotation in infants: Convergent evidence. *Infancy, 19*(1), 103-116. https://doi.org/10.1111/infa.12033

R Core Team (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Rahe, M., & Jansen, P. (2022). Sex differences in mental rotation: the role of stereotyped material, perceived performance and extrinsic spatial ability. *Journal of Cognitive Psychology, 34*(3), 400-409. https://doi.org/10.1080/20445911.2021.2011896

Rahe, M., & Jansen, P. (2023). Does mindfulness help to overcome stereotype threat in mental rotation in younger and older adolescents?. *Psychological Research, 87*(2), 624-635. https://doi.org/10.1007/s00426-022-01666-y

Rahe, M., & Quaiser-Pohl, C. (2020). Cubes or pellets in mental-rotation tests: Effects on gender differences and on the performance in a subsequent math test. *Behavioral Sciences, 10*(1), 8–15. https://doi.org/10.3390/bs10010012

Rahe, M., & Quaiser-Pohl, C. (2023). Can (perceived) mental-rotation performance mediate gender differences in math anxiety in adolescents and young adults?. *Mathematics Education Research Journal, 35*, 255–279. https://doi.org/10.1007/s13394-021-00387-6

Rahe, M., Ruthsatz, V., & Quaiser-Pohl, C. (2021). Influence of the stimulus material on gender differences in a mental-rotation test. *Psychological Research, 85*, 2892-2899. https://doi.org/10.1007/s00426-020-01450-w

Rahe, M., Weigelt, M., & Jansen, P. (2022). Mental rotation with colored cube figures. *Consciousness and Cognition, 102*, 103350. http://doi.org/10.1016/j.concog.2022.103350

Ralph, Y. K., Berinhout, K., & Maguire, M. J. (2021). Gender differences in mothers' spatial language use and children's mental rotation abilities in Preschool and Kindergarten. *Developmental Science, 24*(2), e13037. https://doi.org/10.1111/desc.13037

Rapgay, L., & Bystrisky, A. (2009). Classical mindfulness: an introduction to its theory and practice for clinical application. *Annals of the New York Academy of Sciences, 1172*, 148–162. https://doi.org/10.1111/j.1749-6632.2009.04405.x

Redick, T. S., Calvo, A., Gay, C. E., & Engle, R. W. (2011). Working memory capacity and go/no-go task performance: Selective effects of updating, maintenance, and inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(2), 308–324. https://doi.org/10.1037/a0022216

Repo, S., Elovainio, M., Pyörälä, E., Iriarte-Lüttjohann, M., Tuominen, T., Härkönen, T., ... & Paunio, T. (2022). Comparison of two different mindfulness interventions among health care students in Finland: a randomised controlled trial. *Advances in Health Sciences Education, 27*(3), 709-734. https://doi.org/10.1007/s10459-022-10116-8

Rights, J. D., & Sterba, S. K. (2019). Quantifying explained variance in multilevel models: An integrative framework for defining R-squared measures. *Psychological Methods, 24*(3), 309–338. https://doi.org/10.1037/met0000184

Roberts, J. E., & Bell, M. A. (2003). Two-and three-dimensional mental rotation tasks lead to different parietal laterality for men and women. *International Journal of Psychophysiology, 50*(3), 235-246. https://doi.org/10.1016/S0167-8760(03)00195-8

Robison, M. K., & Brewer, G. A. (2020). Individual differences in working memory capacity and the regulation of arousal. *Attention, Perception, & Psychophysics, 82*, 3273-3290. https://doi.org/10.3758/s13414-020-02077-0

Robison, M. K., & Unsworth, N. (2019). Pupillometry tracks fluctuations in working memory performance. *Attention, Perception, & Psychophysics, 81*, 407-419. https://doi.org/10.3758/s13414-018-1618-4

Roemer, L., Williston, S. K., & Rollins, L. G. (2015). Mindfulness and emotion regulation. *Current Opinion in Psychology, 3,* 52–57. https://doi.org/10.1016/j.copsyc.2015.02.006

Rogers, R. D., & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124*(2), 207–231. https://doi.org/10.1037/0096-3445.124.2.207

Ruthsatz, V., Neuburger, S., Jansen, P., & Quaiser-Pohl, C. (2014). Pellet figures, the feminine answer to cube figures? Influence of stimulus features and rotational axis on the mental-rotation performance of fourth-grade boys and girls. In *Spatial Cognition IX: International Conference, Spatial Cognition 2014, Bremen, Germany, September 15-19, 2014. Proceedings 9* (pp. 370-382). Springer International Publishing. https://doi.org/10.1007/978-3-319-11215-2_26

Ruthsatz, V., Neuburger, S., Jansen, P., & Quaiser-Pohl, C. (2015). Cars or dolls? Influence of the stereotyped nature of the items on children's mental-rotation performance. *Learning and Individual Differences, 43*, 75–82. https://doi.org/10.1016/j.lindif.2015.08.016

Ruthsatz, V., Neuburger, S., Rahe, M., Jansen, P., & Quaiser-Pohl, C. (2017). The gender effect in 3D-Mental-rotation performance with familiar and gender-stereotyped objects–a study with elementary school children. *Journal of Cognitive Psychology, 29*(6), 717–730. https://doi.org/10.1080/20445911.2017.1312689

Sánchez-Sánchez, L. C., Franco, C., Amutio, A., García-Silva, J., & González-Hernández, J. (2023). Influence of Mindfulness on Levels of Impulsiveness, Moods and Pre-Competition Anxiety in Athletes of Different Sports. In *Healthcare* (Vol. 11, No. 6, p. 898). MDPI. https://doi.org/10.3390/healthcare11060898

Schofield, T. P., Creswell, J. D., & Denson, T. F. (2015). Brief mindfulness induction reduces inattentional blindness. *Consciousness and cognition, 37*, 63-70. https://doi.org/10.1016/j.concog.2015.08.007

Schonert-Reichl, K. A., Oberle, E., Lawlor, M. S., Abbott, D., Thomson, K., Oberlander, T. F., & Diamond, A. (2015). Enhancing cognitive and social–emotional development through a simple-to-administer mindfulness-based school program for elementary school children: A randomized controlled trial. *Developmental Psychology, 51*(1), 52–66. https://doi.org/10.1037/a0038454

Schöning, S., Engelien, A., Kugel, H., Schäfer, S., Schiffbauer, H., Zwitserlood, P., et al. (2007). Functional anatomy of visuo-spatial working memory during mental rotation is influenced by sex, menstrual cycle, and sex steroid hormones. *Neuropsychologia, 45*, 3203–3214. https://doi.org/10.1016/j.neuropsychologia.2007.06.011

Schroter, F. A., Siebertz, M., & Jansen, P. (2023). The Impact of a Short Body–Focused Meditation on Body Ownership and Interoceptive Abilities. *Mindfulness, 14*(1), 159-173. https://doi.org/10.1007/s12671-022-02039-7

Seeley, E. A., & Gardner, W. L. (2003). The "selfless" and self-regulation: The role of chronic other-orientation in averting self-regulatory depletion. *Self and Identity, 2*(2), 103-117. https://doi.org/10.1080/15298860309034

Segal, Z. V., Williams, J. M. G., & Teasdale, J. D. (2013). *Mindfulness-based cognitive therapy for depression: a new approach to preventing relapse*. New York, NY: Guilford Press.

Shapiro, L. (2019). *Embodied cognition* (2nd ed.). Routledge/Taylor & Francis Group. https://doi.org/10.4324/9781315180380

Shea, D. L., Lubinski, D., & Benbow, C. P. (2001). Importance of assessing spatial ability in intellectually talented young adolescents: A 20-year longitudinal study. *Journal of Educational Psychology, 93*(3), 604–614. https://doi.org/10.1037/0022-0663.93.3.604

Shepard, R. N., & Cooper, L. A. (1986). *Mental images and their transformations.* The MIT Press.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*(3972), 701–703. https://doi.org/10.1126/science.171.3972.701

Sorby, S. (2009). Developing spatial cognitive skills among middle school students. *Cognitive Processing, 10*(2), 312-315. https://doi.org/10.1007/s10339-009-0310-y

Spencer, S. J., Logel, C., & Davies, P. G. (2016). Stereotype threat. *Annual review of psychology, 67*, 415-437. https://doi.org/10.1146/annurev-psych-073115-103235

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of experimental social psychology, 35*(1), 4-28. https://doi.org/10.1006/jesp.1998.1373

St. Martin, M. W., Vidic, Z., & Oxhandler, R. (2022). Exploring the impact of a mindfulness meditation class on college student stress levels and quality of life. *Journal of American College Health,* 1-7. https://doi.org/10.1080/07448481.2022.2133962

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. Behavior Research Methods, *Instruments, & Computers, 31*(1), 137–149. https://doi.org/10.3758/BF03207704

Stanley, S. (2013). From discourse to awareness: Rhetoric, mindfulness, and a psychology without foundations. *Theory & Psychology, 23*(1), 60-80. https://doi.org/10.1177/0959354312463261

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*(5), 797–811. https://doi.org/10.1037/0022-3514.69.5.797

Steinhauer, S. R., Siegle, G. J., Condray, R., & Pless, M. (2004). Sympathetic and parasympathetic innervation of pupillary dilation during sustained processing. *International Journal of Psychophysiology, 52*(1), 77–86. https://doi.org/10.1016/j.ijpsycho.2003.12.005

Stieff, M. (2007). Mental rotation and diagrammatic reasoning in science. *Learning and Instruction, 17*(2), 219-234. https://doi.org/10.1016/j.learninstnic.2007.01.012

Suess, H. M., Oberauer, K., Wittmann, W. W., Wilhelm, O., & Schulze, R. (2002). Working memory capacity explains reasoning ability – And a little bit more. *Intelligence, 30*, 261–288. https://doi.org/10.1016/S0160-2896(01)00100-3

Tanay, G., & Bernstein, A. (2013). State Mindfulness Scale (SMS): Development and initial validation. Psychological *Assessment, 25*(4), 1286–1299. https://doi.org/10.1037/a0034044

Tang, Y. Y., Ma, Y., Wang, J., Fan, Y., Feng, S., Lu, Q., Yu, Q., Sui, D., Rothbart, M. K., Fan, M., & Posner, M. I. (2007). Short-term meditation training improves attention and self-regulation. *Proceedings of the national Academy of Sciences*, *104*(43), 17152-17156. https://doi.org/10.1073/pnas.0707678104

Tapper, K., Shaw, C., Ilsley, J., Hill, A. J., Bond, F. W., & Moore, L. (2009). Exploratory randomised controlled trial of a mindfulness-based weight loss intervention for women. *Appetite, 52*(2), 396-404. https://doi.org/10.1016/j.appet.2008.11.012

Tartre, L. A., & Fennema, E. (1995). Mathematics achievement and gender: A longitudinal study of selected cognitive and affective variables [grades 6–12]. *Educational studies in mathematics, 28*(3), 199-217. https://doi.org/10.1007/BF01274173

Teasdale, J. D., Moore, R. G., Hayhurst, H., Pope, M., Williams, S., & Segal, Z. V. (2002). Metacognitive awareness and prevention of relapse in depression: Empirical evidence. *Journal of Consulting and Clinical Psychology, 70*(2), 275–287. https://doi.org/10.1037/0022-006X.70.2.275

Teasdale, J. D., Segal, Z., & Williams, J. M. G. (1995). How does cognitive therapy prevent depressive relapse and why should attentional control (mindfulness) training help?. *Behaviour Research and therapy, 33*(1), 25-39. https://doi.org/10.1016/0005-7967(94)E0011-7

Tempel, T., & Neumann, R. (2016). Gender role orientation moderates effects of stereotype activation on test performances. *Social Psychology, 47*(2), 63–73. https://doi.org/10.1027/1864-9335/a000259

Thera, N. (2005). *The heart of buddhist meditation: Satipaṭṭhāna: A handbook of mental training based on the Buddha's way of mindfulness, with an anthology of relevant texts translated from the Pali and Sanskrit.* Buddhist Publication Society.

Thompson, E., & Varela, F. J. (2001). Radical embodiment: Neural dynamics and consciousness. *Trends in Cognitive Sciences, 5*(10), 418–425. https://doi.org/10.1016/S1364-6613(00)01750-2

Titze, C., Jansen, P., & Heil, M. (2011). Single-sex school girls outperform girls attending a co-educative school in mental rotation accuracy. *Sex Roles, 65*, 704-711. https://doi.org/ 704-711. 10.1007/s11199-011-9947-y

Tolkien, J. R. R. (2005). *The Lord of the Rings: Part 1: The Fellowship of the Ring. 50th Anniversary Edition.* HarperCollins UK.

Toth, A. J., & Campbell, M. J. (2021). Reply to:" Concerns about cognitive performance at chance level". *Scientific Reports, 11*(1), 15536. https://doi.org/10.1038/s41598-021-93954-7

Unsworth, N., & Robison, M. K. (2015). Individual differences in the allocation of attention to items in working memory: Evidence from pupillometry. *Psychonomic Bulletin & Review, 22*, 757–765. https://doi.org/10.3758/s13423-014-0747-6

Unsworth, N., & Robison, M. K. (2016). Pupillary correlates of lapses of sustained attention. *Cognitive, Affective, & Behavioral Neuroscience, 16*(4), 601–615. https://doi.org/10.3758/s13415-016-0417-4

Unsworth, N., & Robison, M. K. (2017a). A locus coeruleusnorepinephrine account of individual differences in working memory capacity and attention control. *Psychonomic Bulletin & Review, 24*, 1282–1311. https://doi.org/10.3758/s13423-016-1220-5

Unsworth, N., & Robison, M. K. (2017b). The importance of arousal for variation in working memory capacity and attention control: A latent variable pupillometry study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*, 1962–1987. https://doi.org/10.1037/xlm0000421

Unsworth, N., & Robison, M. K. (2018). Tracking working memory maintenance with pupillometry. *Attention, Perception, & Psychophysics, 80*, 461–484. https://doi.org/10.3758/s13414-017-1455-x

Ussher, M., Spatz, A., Copland, C., Nicolaou, A., Cargill, A., Amini-Tabrizi, N., & McCracken, L. M. (2014). Immediate effects of a brief mindfulness-based body scan on patients with chronic pain. *Journal of behavioral medicine, 37*(1), 127-134. https://doi.org/10.1007/s10865-012-9466-5

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin, 139*(2), 352–402. https://doi.org/10.1037/a0028446

van den Brink, R. L., Murphy, P. R., & Nieuwenhuis, S. (2016). Pupil diameter tracks lapses of attention. *PLoS ONE, 11*(10), e0165274. https://doi.org/10.1371/journal.pone.0165274

van der Wel, P., & van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychonomic bulletin & review, 25*, 2005-2015. https://doi.org/10.3758/s13423-018-1432-y

Vandenberg, S. G., & Kuse, A. R. (1978). Mental Rotations, a Group Test of Three-Dimensional Spatial Visualization. *Perceptual and Motor Skills, 47*(2), 599–604. https://doi.org/10.2466/pms.1978.47.2.599

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: cognitive science and human experience*. Cambridge: MIT Press.

Voyer, D. (1995). Effect of practice on laterality in a mental rotation task. *Brain and Cognition, 29*(3), 326-335. https://doi.org/10.1006/brcg.1995.1285

Voyer, D. (2011). Time limits and gender differences on paper-and-pencil tests of mental rotation: A meta-analysis. *Psychonomic Bulletin and Review, 18*(2), 267–277. https://doi.org/10.3758/s13423-010-0042-0

Voyer, D., Butler, T., Cordero, J., Brake, B., Silbersweig, D., Stern, E., & Imperato-McGinley, J. (2006). The relation between computerized and paper-and-pencil mental rotation tasks: A validation study. *Journal of Clinical and Experimental Neuropsychology, 28*(6), 928–939. https://doi.org/10.1080/13803390591004310

Voyer, D., & Jansen, P. (2016). Sex differences in chronometric mental rotation with human bodies. *Psychological Research, 80*(6), 974–984. https://doi.org/10.1007/s00426-015-0701-x

Voyer, D., Jansen, P., & Kaltner, S. (2017). Mental rotation with egocentric and object-based transformations. *Quarterly Journal of Experimental Psychology, 70*(11), 2319–2330. https://doi.org/10.1080/17470218.2016.1233571

Voyer, D., Nolan, C., & Voyer, S. (2000). The relation between experience and spatial performance in men and women. *Sex Roles, 43*(11–12), 891–915. https://doi.org/https://doi.org/10.1023/A:1011041006679

Voyer, D., Saint-Aubin, J., Altman, K., & Doyle, R. A. (2020). Sex differences in tests of mental rotation: Direct manipulation of strategies with eye-tracking. *Journal of Experimental Psychology: Human Perception and Performance, 46*(9), 871–889. https://doi.org/10.1037/xhp0000752

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin, 117*(2), 250–270. https://doi.org/10.1037/0033-2909.117.2.250

Vveinhardt, J., & Kaspare, M. (2022). The Relationship between Mindfulness Practices and the Psychological State and Performance of Kyokushin Karate Athletes. *International journal of environmental research and public health, 19*(7), 4001. https://doi.org/10.3390/ijerph19074001

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial Ability for STEM Domains: Aligning Over 50 Years of Cumulative Psychological Knowledge Solidifies Its Importance. *Journal of Educational Psychology, 101*(4), 817–835. https://doi.org/10.1037/a0016127

Wai, J., Lubinski, D., Benbow, C. P., & Steiger, J. H. (2010). Accomplishment in Science, Technology, Engineering, and Mathematics (STEM) and Its Relation to STEM Educational Dose: A 25-Year Longitudinal Study. *Journal of Educational Psychology, 102*(4), 860–871. https://doi.org/10.1037/a0019454

Wallace, B. A. (1999). The Buddhist tradition of Samatha: methods for refining and examining consciousness. *Journal of Consciousness Studies, 6*(2–3), 175–187.

Wallace, B. A. (2006). *The attention revolution: Unlocking the power of the focused mind*. Boston, MA: Wisdom Publications.

Walton, G. M., & Cohen, G. L. (2003). Stereotype Lift. *Journal of Experimental Social Psychology, 39*(5), 456–467. https://doi.org/10.1016/S0022-1031(03)00019-2

Waszak, F., Hommel, B., & Allport, A. (2003). Task-switching and long-term priming: Role of episodic stimulus–task bindings in task-shift costs. *Cognitive psychology, 46*(4), 361-413. https://doi.org/10.1016/s0010-0285(02)00520-0

Weger, U. W., Hooper, N., Meier, B. P., & Hopthrow, T. (2012). Mindful maths: Reducing the impact of stereotype threat through a mindfulness exercise. *Consciousness and cognition, 21*(1), 471-475. https://doi.org/10.1016/j.concog.2011.10.011

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta psychologica, 41*(1), 67-85. https://doi.org/10.1016/0001-6918(77)90012-9

Wickens, C., Kramer, A., Vanasse, L., & Donchin, E. (1983). Performance of concurrent tasks: a psychophysiological analysis of the reciprocity of information-processing resources. *Science, 221*(4615), 1080-1082. https://doi.org/10.1126/science.6879207

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. ISBN 978-3-319-24277-4. https://ggplot2.tidyverse.org

Williams, J. M. G., & Kabat-Zinn, J. (2011). Mindfulness: diverse perspectives on its meaning, origins, and multiple applications at the intersection of science and dharma. *Contemporary Buddhism, 12*, 1–18. https://doi.org/10.1080/14639947.2011.564811

Wilson, A. D., & Golonka, S. (2013). Embodied cognition is not what you think it is. *Frontiers in psychology, 4*, 58. https://doi.org/10.3389/fpsyg.2013.00058

Winter, B., & Wieling, M. (2016). How to analyze linguistic change using mixed models, Growth Curve Analysis and Generalized Additive Modeling. *Journal of Language Evolution, 1*(1), 7–18. https://doi.org/10.1093/jole/lzv003

Witkin, H. A. (1971). *A manual for the embedded figures tests*. Consulting Psychologists Press.

Witkin, H. A., & Asch, S. E. (1948). Studies in space orientation. IV. Further experiments on perception of the upright with displaced visual fields. *Journal of Experimental Psychology, 38*(6), 762–782. https://doi.org/10.1037/h0053671

Wolkin, J. R. (2015). Cultivating multiple aspects of attention through mindfulness meditation accounts for psychological well-being through decreased rumination. *Psychology research and behavior management, 8*, 171-180. https://doi.org/10.2147/PRBM.S31458

Wu, H.-K., & Shah, P. (2004). Exploring visuospatial thinking in chemistry learning. *Science Education, 88*(3), 465-492. https://doi.org/10.1002/sce.l0126

Xie, F., Zhang, L., Chen, X., & Xin, Z. (2020). Is Spatial Ability Related to Mathematical Ability: a Meta-analysis. *Educational Psychology Review, 32*(1), 113–155. https://doi.org/10.1007/s10648-019-09496-y

Yang, J., Wu, D., Luo, J., Xie, S., Chang, C., & Li, H. (2020). Neural correlates of mental rotation in preschoolers with high or low working memory capacity: An fNIRS study. *Frontiers in Psychology, 11*, 568382. https://doi.org/10.3389/fpsyg.2020.568382

Yilmaz, H. B. (2017). On the development and measurement of spatial ability. *International Electronic Journal of Elementary Education, 1*(2), 83–96. Retrieved from https://www.iejee.com/index.php/IEJEE/article/view/279

Yoshitomi, T., Ito, Y., & Inomata, H. (1985). Adrenergic excitatory and cholinergic inhibitory innervations in the human iris dilator. *Experimental eye research, 40*(3), 453-459. https://doi.org/10.1016/0014-4835(85)90158-7

Zacks, J. M., Mires, J., Tversky, B., & Hazeltine, E. (2000). Mental spatial transformations of objects and perspective. *Spatial Cognition and Computation, 2*(4), 315–332. https://doi.org/10.1023/A:1015584100204

Zacks, J. M., & Tversky, B. (2005). Multiple systems for spatial imagery: Transformations of objects and bodies. *Spatial Cognition and Computation, 5*(4), 271-306. https://doi.org/10.1207/s15427633scc0504_1

Zeidan, F., Gordon, N. S., Merchant, J., & Goolkasian, P. (2010). The effects of brief mindfulness meditation training on experimentally induced pain. *The Journal of Pain, 11*(3), 199-209. https://doi.org/10.1016/j.jpain.2009.07.015

Zeidan, F., Johnson, S. K., Diamond, B. J., David, Z., & Goolkasian, P. (2010). Mindfulness meditation improves cognition: Evidence of brief mental training. *Consciousness and cognition, 19*(2), 597-605. https://doi.org/10.1016/j.concog.2010.03.014

Zoccolotti, P., Antonucci, G., & Spinelli, D. (1993). The gap between rod and frame influences the rod-and-frame effect with small and large inducing displays. *Perception & Psychophysics, 54*, 14-19. https://doi.org/10.3758/BF03206933

Zokaei, N., Board, A. G., Manohar, S. G., & Nobre, A. C. (2019). Modulation of the pupillary response by the content of visual working memory. *Proceedings of the National Academy of Sciences, 116*(45), 22802–22810. https://doi.org/10.1073/pnas.1909959116

Zuur, A. F., Ieno, E. N., & Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in ecology and evolution, 1*(1), 3-14. https://doi.org/10.1111/j.2041-210X.2009.00001.x

Eidesstattliche Versicherung

Robert Bauer

Geburtsdatum: 13.01.1989

Ich erkläre hiermit an Eides Statt, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Textpassagen, Daten, Bilder oder Grafiken sind unter Angabe der Quelle gekennzeichnet.

Bei der Auswahl und Auswertung folgenden Materials haben mir die nachstehend aufgeführten Personen in der jeweils beschriebenen Weise unentgeltlich geholfen, dies ist auch in der Dissertation an den entsprechenden Stellen explizit ausgewiesen (anhand von Zitationen gemeinsamer Publikationen in Fußnoten sowie Sektion 8.3):

1. Petra Jansen (Promotionsbetreuung)
2. Bianca Günther (Co-Autorin)
3. Leonardo Jost (Co-Autor)

Weitere Personen waren an der inhaltlich-materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs- beziehungsweise Beratungsdiensten (Promotionsberater oder anderer Personen) in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Ich versichere an Eides Statt, dass ich nach bestem Wissen die reine Wahrheit gesagt und nichts verschwiegen habe.

Vor Aufnahme der obigen Versicherung an Eides Statt wurde ich über die Bedeutung der eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung belehrt.

……………………. ……………………………………………

Regensburg Unterschrift des Promovenden