An investigation into the perceptual influence of prior information in the human cognitive system

Dall, Jonas Olsen

DOI (link to publication from Publisher):
10.54337/aau470867787

Publication date:
2021

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

AN INVESTIGATION INTO THE PERCEPTUAL INFLUENCE OF PRIOR INFORMATION IN THE HUMAN COGNITIVE SYSTEM

BY
JONAS OLSEN DALL

DISSERTATION SUBMITTED 2021

AALBORG UNIVERSITY
DENMARK

AN INVESTIGATION INTO THE PERCEPTUAL INFLUENCE OF PRIOR INFORMATION IN THE HUMAN COGNITIVE SYSTEM

by

Jonas Olsen Dall

AALBORG UNIVERSITY
DENMARK

Dissertation submitted

.

Dissertation submitted:	23/12/2021
PhD supervisor:	Associate Prof. Thomas Alrik Sørensen, Aalborg University
Assistant PhD supervisor:	Associate Prof. Raymond C. K. Chan, Chinese Academy of Sciences
PhD committee:	Associate Professor Laura Petrini Aalborg University (chair)
	Assistant Professor Iris Wiegand Radboud University
	Professor Berit Brogaard University of Miami
PhD Series:	Faculty of Humanities, Aalborg University

ISSN (online): 2246-123X
ISBN (online): 978-87-7573-954-7

Published by:
Aalborg University Press
Kroghstræde 3
DK – 9220 Aalborg Ø
Phone: +45 99407140
aauf@forlag.aau.dk
forlag.aau.dk

Printed in Denmark by Rosendahls, 2022

ENGLISH SUMMARY

Classic studies in psychology have shown that attention span and memory capacity is highly limited. These capacities further decrease if the to be remembered stimuli are something unfamiliar. The aim of this thesis is therefore to study the mechanics behind this change in capacity.

Study 1 had the dual purpose of both studying these phenomena and showing that stimuli specific expertise was possible to build in adulthood. My co-authors and I did this by extending a previous study (Sørensen & Kyllingsbæk, 2012) from looking at Visual Short-Term Memory (VSTM) changes in early life related Latin letter expertise to late-in-life acquired Japanese hiragana expertise. We did this by studying the Visual Short-Term Memory capacity (*K*) for the comparisons (Japanese hiragana) and two control stimuli (Pictures and Latin letters) in participants at three different levels of stimuli specific expertise: Danish university students with no Japanese expertise (Novice), Danish students studying Japanese at the University (Trained), and Native Japanese University students (Expert). We found that there were no significant differences between the groups in the two control stimuli. There was, however, an expertise related influence in the hiragana condition (Novice < Trained < Expert). Furthermore, despite the more visually complex Japanese hiragana, there was no significant difference in the native languages. Furthermore, no differences were found for Latin for Novice and Trained compared to hiragana for the Japanese, despite the fact that Japanese hiragana is visually more complex.

Study 2 studied the relationship between this apparent visual complexity and stimuli specific expertise. This was achieved by using Chinese characters as they have an approximation for both stroke count as visual complexity and frequency of use as stimuli specific expertise. Stimuli were divided into four conditions based on stroke count and frequency of use: High-Stroke High-Frequency, High-Stroke Low-Frequency, Low-Stroke High-Frequency, and Low-Stroke Low-Frequency. The study found no significant influence of stroke count while finding that frequency of use influenced both *K* and processing speed (*C*) without influencing threshold for perception (t_0). This raises the question of whether *K* and *C* have an effect on stimuli specific expertise or if the differences are due to a stronger mental resolution of stimuli stored in VSTM.

This question of stimuli specific expertise and mental resolution were studied in study 3 by using continuous categories. Stimuli were divided into two conditions: colour as continuous categories and Latin letters as a control condition. Participants used a colour wheel to report the stimuli on the colour trial. The precision of these answers were based on the distance between answers on the colour wheel to what they were shown. Expertise was modulated by dividing participants into two groups: Danish psychology students with no training in colour, design, or other visual arts (Novice)

as well as Danish students at School of Visual Arts at the Royal Danish Academy of Fine Arts (Expert). We did not find any difference between the two groups in either colour precision, *C*, or t_0. However, colour experts showed a significantly higher *K* for colours than Novices.

The studies comprising my thesis may suggest that stimuli specific expertise can be gained late in life and increases both memory capacity and processing speed. These changes can furthermore not be explained by increases in mental resolution of the stored items but instead as increases in stimuli specific memory capacity.

CHINESE SUMMARY

心理学的经典研究表明，注意力广度和记忆的容量都是十分有限的。如果需要记住的刺激并非熟知领域，那么这些能力也会进一步降低。**因此，本篇**论文的目的是研究这种容量变化背后的运作方式。

研究 1 具有双重目的，首先要研究这些现象，其次**要**证明人在成年期建立特定专业知识刺激是可能的。我和我的合著者们扩展了之前的**一**项研究（Sørensen & Kyllingsbæk, 2012 **年**），**我**们从研究青年时就掌握拉丁字母相关知识扩展到晚年掌握日语平假名专业知识的视觉短期记忆(VSTM)**的**变化来着手研究。我们通过研究视觉短期记容量K**来比**较（日语平假名）和两种控制刺激（图片和拉丁字母），试验参与者们拥有三种特定专业刺激：没有日语专业知识的丹麦大学生（新手），在大学学习日语的丹麦学生（受训过）和日本本土大学生（专家）。我们发现在两种控制刺激下，两组之间没有显着差异。然而却存在平假名专业知识相关的影响（新手 < **受**训 < 专家）。此外，初学者和受训者在拉丁语没有也未出现差异，尽管日本平假名在视觉上通常更加复杂。

研究 2 探究了这种明显的视觉复杂性与特定专业知识刺激之间的关系。这是通过使用汉字来实现的，因为它们作为视觉复杂性的笔画计数和刺激特定专业知识的使用频率近似。刺激根据行程计数和使用频率分为四种条件：高行程高频**、高行程低**频、低行程高频和低行程低频。该研究发现中风计数没有显着影响，同时发现使用频率会影响K **和** 进展速度C，**然而却不影响感知**阈值 t_0。该发现引出了一个问题，即特定专业知识刺激是否真正影响K **和** C，**或者**说差异是否由于储存在 VSTM **中的刺激的心理分**辨**率更**强。

研究 3 是通过使用连续类别来研究刺激特定专业知识和分辨率的问题。刺激分为两种条件：着色为连续类别和拉丁字母作为控制条件。参与者使用色轮报告颜色试验的刺激。答案的精确度取决于色轮上的答案与显示内容之间的距离。试验通过专业程度将参与者分为两组：没有接受过色彩、设计或其他视觉艺术培训的丹麦心理学学生（新手）以及丹麦皇家美术学院视觉艺术学院的丹麦学

生（专家）。我们没有发现两组在颜色精度、C **或** t_0 **方面有任何差异。然而根据**试验结果显示，色彩专家的色彩K 值明显高于新手。

一系列的研究包括此篇论文可能表明，刺激特定专业知识，并提高记忆容量和进展速度可在晚年获取。此外，这些变化不能用存储项目的心理分辨率的增加来解释，而是通过刺激特定记忆容量的增加来说明。

DANSK RESUME

Klassiske psykologiske studier har fundet, at opmærksomheden og hukommelseskapaciteten er meget afgrænset, og er afgrænset yderligere, hvis det er noget atypisk man skal huske. Målet med denne afhandling er derfor at undersøge mekanismerne bag denne ændring.

Studie 1 har den dobbelte hensigt at studere det ovenstående fænomen, og at vise det er muligt at opbygge stimuli-specifik ekspertise som voksen. Mine medforfattere og jeg gjorde dette ved at udvide et tidligere studie (Sørensen & Kyllingsbæk, 2012) der undersøgte den visuelle korttidshukommelsen fra aldersrelateret ekspertise med latinske bogstaver, til voksen erhvervet japansk hiragana-ekspertise. Vi gjorde dette ved at studere den visuelle korttidshukommelseskapacitet (*K*) for den kritiske (japansk hiragana) og to kontrolkonditioner (billeder og latinske bogstaver) hos forsøgspersoner med tre forskellige ekspertise-niveauer: danske universitetsstuderende uden nogen japansk ekspertise (Novice), danske universitetsstuderende der studerer japansk på universitetet (Trained), og universitetsstuderende med japansk som deres modersmål (Expert). Vi fandt ingen signifikant forskel mellem grupperne for de to kontrol-stimuli, men fandt en ekspertise-relateret forskel i hiragana-konditionen (Novice < Trained < Expert). Der var yderligere ingen forskel mellem modersmålene, på trods af hiragana generelt er mere visuelt kompleks.

Studie 2 undersøgte forskellen mellem den visuelle kompleksitet og stimuli-specifik ekspertise. Vi gjorde dette ved at sammenligne kinesiske tegn, da de kan deles op i, antallet af streger, og hvor ofte de bliver brugt, hvilket giver et estimat for visuel kompleksitet og stimuli-specifik ekspertise henholdsvis. Vi delte stimuli op i fire konditioner på baggrund af antal streger og hvor ofte tegnene blev brugt: High-Stroke High-Frequency, High-Stroke Low-Frequency, Low-Stroke High-Frequency, and Low-Stroke Low-Frequency. Antallet af streger havde ingen signifikant indflydelse på hverken *K*, forarbejdningshastighed (*C*), eller tærsklen for opmærksomhed (t_0), hvorimod hvor ofte tegnet blev brugt havde en signifikant indflydelse på *K* og *C*. Dette rejser spørgsmålet om hvorvidt denne indflydelse af stimuli-specifik ekspertise rent faktisk påvirker *K* og *C,* eller om disse forskelle i stedet er et udtryk for en øget mental opløsning af det stimuli, der er gemt i VSTM.

Dette spørgsmål om stimuli-specifik ekspertise og mental opløsning blev undersøgt i studie 3, ved brug af kontinuerlige kategorier. Vi delte derfor stimuli ind i to konditioner: farver som den kontinuerlige kategori og latinske bogstaver som en kontrolkondition. Forsøgspersonerne brugte et farvehjul til at rapportere hvilke farver de så i farvekonditionen. Hvor længe den valgte farve på farvehjulet var, blev brugt som et mål for præcision. Ekspertise blev målt ved at dele forsøgspersonerne ind i to grupper: Danske psykologistuderene, der ikke havde nogen træning med farver, design, eller andre visuelle kunstformer (Novice), og danske studerende på Billedkunstskolen hos det Kongelige Danske Kunstakademi (Expert). Vi fandt ingen forskel mellem de to grupper i hverken farvepræcision, *C,* eller t_0, dog havde Experts højere *K* for farver end Novices.

Studierne, der danner grundlag for afhandlingen, indikerer at stimuli-specifik ekspertise kan erhverves sent i livet og forbedre både hukommelseskapaciteten og processeringshastighed. Disse ændringer kan ligeledes ikke forklares via forbedring i den mentale opløsning, men i stedet via en stigning i stimuli-specifik hukommelseskapacitet.

ACKNOWLEDGEMENTS

I would like to thank everyone who have supported me through the writing of my PhD. I would especially like to thank my Chinese supervisor Raymond C.K. Chan and his lab with a special thanks to my co-authors Wang, Y. M. and Cai, X. L. for their help recruiting participants and collecting data. I would also like to thank Katsumi Watanabe for welcoming me to his lab and allowing me to collect data when I was still a master's student.

I would like to extend a special thanks to my colleagues at Centre for Cognitive Neuroscience in Aalborg and my Danish supervisor Thomas Alrik Sørensen. I would like to thank my family and friends who have listen to me practising more than one lecture on them. Finally, I would like to thank my proof-readers for proof-read an early draft of this thesis.

DEFINITIONS

Natural training: Training, not as a part of an experiment, but for other reasons (e.g. learning a new language as part of one's education or natural language acquisition).

Sigmoid function: Is given by the formula $f(x) = \frac{1}{1+e^{-x}}$ and is characterised by an 'S' curve, which represents a gradual exponential increase until the increase starts to exponentially diminish, until the curve reaches an asymptote (Das & Cakmak, 2018).

Hazard function: A hazard function calculates the probability that an event will occur if it has not already happened: $p(X = t | X \geq t)$. Meaning the probability (p) of an event occurring at time *t* ($X = t$), as long as the event has not yet happened ($X \geq t$). This is calculated by dividing the probability function (the probability that the event happens at time 1) by the survival function (the probability that the event has not occurred before *t*) (Everitt & Palmer, 2005).

Poisson distribution: The Poisson distribution is a probability distribution. It indicates the probability at different independent values, with the probability of an event given as $(x) = \frac{\lambda^x}{x!} e^{-\lambda}$. The probability of something happening after x number of occurrences is therefore e to the power of negative λ (the average number of occurrences of a specific event), times λ to the power of x (the number of occurrences), divided by the factorial of x. This roughly means that the higher the value of λ, the more the curve resembles a normal distribution (Everitt & Palmer, 2005).

Perceptual units: Perceptual units are Gestalt-like objects that are groupings of sensory information. This grouping can happen both through space (e.g. black pixels on a white screen creating: L) or time (e.g. letters produced sequentially creating a word). It should however also be noted that the grouping of stimuli into perceptual units does not mean that the stimuli have been processed, just that it is treated as a single unit rather than several unrelated units (Kahneman, 1973).

LIST OF ABBREVIATIONS

STM: Short-Term Memory

VSTM: Visual Short-Term Memory

WM: Working Memory

d: Cohen's *d*

K: Visual short-term memory capacity

C: Processing speed

t: Time

t_0: Threshold for minimal perception

η: Sensory evidence for a specific category

β: Perceptual bias

π: Pertinence value

TABLE OF CONTENTS

TABLE OF FIGURES

CHAPTER 1. INTRODUCTION

The year 1956 have been described as 'a good year' in American experimental psychology with several important papers and conferences which helped spark the cognitive revolution (Miller, 2003). Among these field defining events were the publication of a last-minute lecture by Miller at the 1955 Eastern Psychological Association conference which he had to be cajoled into (Cowan, 2015). Miller's paper (1956) titled "*The magical number seven, plus or minus two: Some limits on our capacity for processing information*", argued that there was seemingly a limit to how many chunks of information could be maintained in short-term memory (STM), with a rough estimate of 7±2. This approximation of 7±2 can still be seen in some textbooks (e.g. Purves et al., 2018) and in studies using memory spans that allow memory strategies (see Cowan, 2001). Later research argued the unsupported visual short-term memory (VSTM) capacity to be approximately 3 – 4 items (e.g. Bundesen, 1990; Cowan, 2001; Sperling, 1960). The difficulty in finding the correct number of memory slots in STM is explained in Miller's (1956) second contribution: Our ability to chunk information. 'Chunking' refers to the ability of "...organizing or grouping the input into familiar units or chunks, and a great deal of learning has gone into the formation of these familiar units" (Miller, 1956). Miller (1956) argues that information stored in the STM is not typically stored in the smallest *bits* of information, but rather as a combination of several components. He exemplifies this by arguing that it is unlikely that five three-letter words are stored as 15 phonemes, and more likely that they are stored as five separate words. Chunking therefore makes it difficult for researchers who are trying to find the actual STM capacity, as they need to control which chunking strategy participants perform to get accurate measures (e.g. Bundesen, 1990; Conway, Kane, Bunting, Hambrik, Wilhelm, Engel, 2005). Some participants, for example, have been able to remember 80 digits after two months of training (Ericsson & Chase, 1982). A number of researchers have developed several methods to prevent participants from creating new chunks of target stimuli, ranging from task demand (e.g. Baddeley, Lewis, & Vallar, 1984; Conway et al., 2005; Daneman & Carpenter, 1980; Peterson, 1969; Richardson & Baddeley, 1975), to showing stimuli at a rate that is too fast for participants to sample enough stimuli to create new chunks (e.g. Bundesen, 1990; Luck, & Vogel, 1997; Sperling, 1960). These methods have meant that debates regarding memory capacity have become more integrated into attention research, primarily because the attention paradigms created by Sperling (1960) have become a common way of circumventing chunking (Cowan, 2001).

The memory capacity not mediated by chunking is highly influential as it has been shown to be a major determinant of cognitive development in childhood (Bayliss, Jarrold, Gunn, & Baddeley, 2003), in old age (Park et al., 2002; Salthouse, 1994), and in individual differences in intellectual abilities (Conway, Kane, & Engle, 2003; Jarrold & Towse, 2006). This make problems in the precision of the measurements even more important. Several studies (e.g. Alvarez & Cavanagh, 2004; Jackson et al., 2014; Oberauer & Eichenberger, 2013; Swan et al., 2016) have shown that this unchunked memory capacity is influenced by several other factors. Among these are stimuli specific expertise and prior information (Dall, et al., 2021; Popov & Reder, 2020; Sun et al., 2011; Sørensen & Kyllingsbæk, 2012). There is however little information on how these factors unchunked memory interacts with each other. This thesis therefore examines the relationship between unchunked memory and stimuli specific expertise and will use VSTM as the 3 – 4 measurement, as it is the suspected limit for the visual sensory system without support from other systems (Bundesen, 1990).. The thesis first provides an introduction to the current theories on how memory capacity is influenced by expertise, as well as discussion on different theories of how to understand the perceived limit on visual STM's capacity (Chapter 2), with special weight given to the theory of visual attention (Chapter 2.3.1). This will lead to an analysis of the methodology used in the experiments of this paper (Chapter 3) followed by a summary of the papers (Chapter 4). The discussion examines how the papers fit into current research and new insights they provide (Chapter 5), followed by concluding remarks (Chapter 6).

CHAPTER 2. ATTENTION

Attention research has a long history in science, from the early observations of giants like Wundt, James, and Cattell (e.g. Jacobs, 1887; James, 1890; Wundt, 1899;), to the cognitive revolution (e.g. Atkinsson & Shiffrin, 1968; Baddeley & Hitch, 1974; Broadbent, 1987; Miller, 1956; Sperling, 1960), and all the way to contemporary research (e.g. Bundesen, 1990; Cowan, 2001; Luck & Vogel, 1997; Wilken & Ma, 2004; Zhang & Luck, 2008). Attention research is split into several research focuses. Among them are how fast stimuli is processed (see e.g. Cattell, 1890; Donders, 1969; Pelli, et al., 2006), how much is actively processed (see e.g. Cowen, 2001; Miller, 1956), the neural trajectory of processing (see e.g. Hubel & Wiesel, 1959, Usher & Cohen, 1999), and when stimuli are consciously processed (see e.g. Ramsøy & Overgaard, 2004).

Processing speed has long been an area of interest for attention research (Cattell, 1890; Donders, 1969). Some researchers have suggested that processing speed is the main limiting factor in VSTM capacity (e.g. Alvarez & Cavanagh, 2004; Cavanagh, 1972; Hulme, Maughan, & Brown, 1991; Hulme, Roodenrys, Brown & Mercer, 1995). Pelli and colleagues (2006) showed that the speed at which a stimulus was identified was influenced by both its visual complexity and by the participant's familiarity with it. Several authors have in that spirit demonstrated that while several thousand trials of training can improve processing speed, there is little effect shown in subsequent memory tests (e.g. Chen, Eng, & Jiang 2006; Olson & Jiang, 2004; Pelli, et al., 2006). Which seems to suggest that the processing speed is not the only limiting factor in VSTM capacity. Level of expertise has, however, been shown to influence processing speed in specific parts of the visual field (Schubert, Finke, Redel, Kluckow, Müller, & Strobach, 2015) which might suggest that visual processing is further divided into smaller local processing speeds.

Hubel and Wiesel showed that early processing was connected to neuronal activity in a series of cat studies (1959, 1962, 1963, 1965; Hubel, 1959). They found that neurons in V1 showed activation preference for specific placement and rotation of lines. This has led to several studies which tried to follow signals from early (e.g. Atick & Redlich, 1990; Hubel & Wiesel, 1959, 1962, 1963, 1965; Hubel, 1959; Marr, 1976) to late processing (e.g. Chadwick, Anjum, Kumaran, Schacter, Spiers, & Hassabis, 2016; Vukovic, Feurra, Shpektor, Myachykov, & Shtyrov, 2017). This have lead researchers to examine similar activation preference but for further down the processing stream. The relationship between faces and specific neural regions has in this regard received a great deal of attention (Kanwisher, McDermott, & Chun, 1997; Gauthier, Skudlarski, Gore, & Anderson, 2000), partly because faces are visually complex stimuli that are efficiently processed (Gauthier & Tarr, 1997) and partly due to their association with specific neural regions (Gauthier, Tarr, Anderson, Skudlarski, 1999; Grill-Spector, Knouf, & Kanwisher, 2004; Tong, Nakayama, Moscovitch, Weinrib, & Kanwisher, 2000).

Faces were some of the earliest objects to be associated with the fusiform gyrus (FFG), which was initially proposed as an area dedicated to face processing (Gross, Rocha-Miranda, & Bender, 1972; Kanwisher, McDermott, & Chun, 1997; Ojemann, Ojemann, & Lettich, 1992). Later studies have shown that high expertise stimuli also elicit a reaction in FFG (e.g. Gauthier, Skudlarski, Gore, & Anderson, 2000; Grelotti, Klin, Gauthier, Skudlarski, Cohen, Gore, Volkmar, & Schultz, 2005; McGugin, Newton, Gore, & Gauthier, 2014). Similar neuronal connections have been found in late processing areas of the brain for both categorical representation (Chadwick, Anjum, Kumaran, Schacter, Spiers, & Hassabis, 2016) and understanding (Vukovic, Feurra, Shpektor, Myachykov, & Shtyrov, 2017). This would suggest that there is some overlap between neurons, specific stimuli, and conceptual categories similar to line rotation in early perception, however McGugin and colleagues (2014) demonstrated FFG activation for non-related items that were visually similar to high expertise items (e.g. sofas and cars). Several authors have built on Hubel and Wiesel's original studies and suggest that attention (and by extension VSTM) is the online activation of neurons' representation in mental templates of the attended stimuli (e.g. Usher & Cohen, 1999).

2.1. MODULATION OF ATTENTION

The topic of memory capacity has within periods been hotly debated with different focuses. These topics range from the memory span limit (Jacobs, 1887), the number of 'slots' in VSTM, the nature of 'chunks' (Miller, 1956), measuring a VSTM unaffected by chunking (Cowan, 2001), as well as what constitutes a 'slot' and how it is manipulated (Scolari, Vogel, & Awh, 2008; Wilken & Ma, 2004; Zhang & Luck, 2008). The current debate was sparked when Luck and Vogel (1997) found that the number of stimulus features (from 1-4) did not change the number of stimuli reported, indicating that several features could be nestled into an object stored in VSTM.

Several authors attempted (Alvarez & Cavanagh, 2004; Wheeler & Treisman, 2002), and failed, to replicate and extend the finding that items in VSTM were unaffected by stimulus complexity. Alvarez and Cavanagh (2004) paved the way for a series of new studies, when they found that visual complexity, contrary to the findings of Luck and Vogel (1997), severely decreased VSTM capacity. Alvarez and Cavanagh (2004) argued that this discrepancy was likely due to the simplistic stimuli used by Luck and Vogel (1997). Alvarez and Cavanagh (2004) further showed a correlation between search rate and memory capacity, and found that items like random polygon and Chinese characters were more difficult to remember than shaded cubes and Latin letters. They argued that these differences were based on the items' visual complexity, which would therefore mean that Chinese characters were inherently more difficult to remember than Latin letters. This was subsequently examined by several authors, who found that native Chinese readers were better at remembering Chinese characters, compared to participants naïve to Chinese (Dall, Wang, Cai, Chan, & Sørensen, 2021; Jackson & Raymond, 2005; Sun, Zimmer & Fu, 2011; Zimmer, Popp, Reith & Krick, 2012). This would suggest that VSTM capacity is influenced by the level of stimuli specific expertise.

Jackson and Raymond (2005) extended the conclusions from Alvarez and Cavanagh (2004) by suggesting that complexity should not be understood on the sole basis of the stimulus, but also the expertise of the participant looking at the stimulus. Jackson and Raymond (2005) divided the complexity of any viewed stimulus into physical and perceived complexity. Physical complexity can be understood as the number of details on the object, whereas perceived complexity is related to the level of expertise the participant has. It should be noted that visual complexity depends on the unique difference between stimuli (Swan, Collins, & Wyble, 2016) and their confusability (Jackson et al., 2015). Perceived complexity would therefore be a combination between the physical complexity mediated by the level of stimulus specific expertise.

This connection between stimuli familiarity and visual complexity has been examined in multiple investigations (e.g. Chen, Eng, & Jiang, 2006; Dall, Katsumi, Sørensen, 2016; Dall et al., 2021; Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999; Grelotti, et al., 2005; James & James, 2013; McGugin, Newton, Gore, & Gauthier, 2014; Olson & Jiang, 2004; Pelli, et al., 2006; Price, et al., 2016; Rossion, Kung, & Tarr, 2004; Sun, Zimmer & Fu 2011; Sørensen & Kyllingsbæk, 2012; Williams & Simons, 2000; Xie & Zhang, 2017a; 2017b; Zimmer, Popp, Reith & Krick 2012). Several of these have tried to evidence a link between attention and expertise by training participants in unfamiliar stimuli (e.g. Chen, Eng, & Jiang, 2006; Olson & Jiang, 2004; Pelli, et al., 2006; Price, et al., 2016; Williams & Simons, 2000). A classic tool amongst these training studies is the use of greebles to examine both the behaviour and neural changes following training with specific stimuli (e.g. Gauthier, Behrmann, & Tarr, 2004; Tarr & Gauthier, 2000). Greebles are small, similar figures made by Scott Yu (Gauthier & Tarr, 1997) that vary in either gender or family (see fig. 1). Several studies have found that it is possible to improve participants' processing speed of unfamiliar writing systems (e.g. Pelli, et al., 2006; Price, et al., 2016), greebles (Gauthier & Tarr, 1997) and geometric figures (e.g. Williams & Simons, 2000), while showing no improvements in VSTM capacity. This coincides with results from neuroimaging studies, which found that high expertise stimuli showed activation in FFG (e.g. Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999; Grelotti, et al., 2005; James & James, 2013; McGugin, et al., 2014; Rossion, Kung, & Tarr, 2004). However, there is still some evidence that makes the connection between FFG and expertise debatable. A study by McGugin and colleagues (2014) reported that while there was an increase in FFG activation when car experts looked at cars, there was likewise an increased activation when looking at structurally similar items, such as sofas. This was despite the car experts showing no significant advantage in discriminating between sofas compared to other non-car stimuli. This expertise-focused line of research suggests that unfamiliar complex stimuli decrease the VSTM capacity, leading to the question of what happens to VSTM in these cases, and where do the VSTM ‘slots’ at lower levels of expertise disappear to?

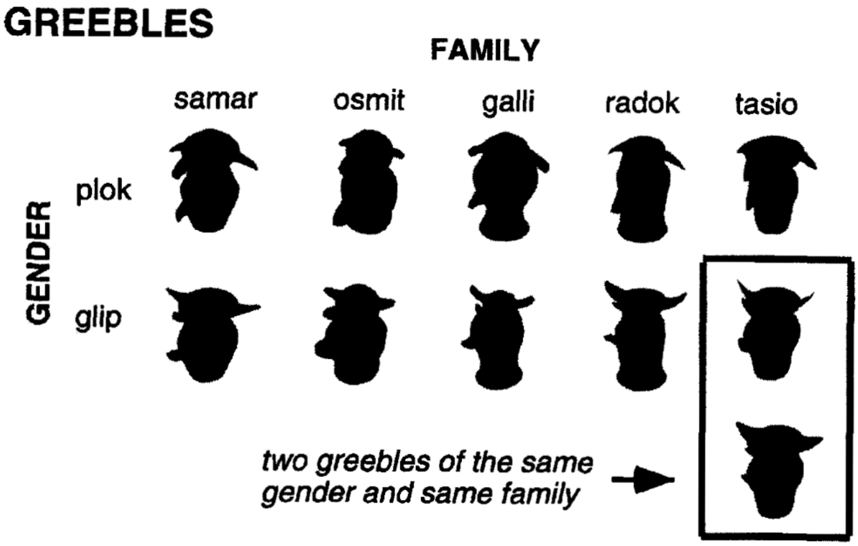

Fig. 1 Shows examples of the different greeble genders (Plok & Glip) and families (Samar, Osmit, Galli, Radok, & Tasio). As well as an example of two greebles from the same gender and family (Gauthier & Tarr, 1997).

The use of resolution paradigms changed the debate by claiming to have found these missing memory slots. Resolution paradigms typically use discrete categories stimuli, where the precision of the answer can be rated (Suchow, Brady, Fougnie, & Alvarez, 2013; Wilken & Ma, 2004; Zhang & Luck, 2008). This means that participants are shown stimuli and thereafter need to reproduce one or more using items such as colour wheels, continuous face space, shapes, or similar to answer (see e.g. Li, Liang, Lee, & Barense, 2019; Lorenc, Pratte, Angeloni, & Tong, 2014; Wilken & Ma, 2004; Zhou, Mondloch, & Emrich, 2018). This means that it is possible to calculate how close participants were to the correct answer, a task that is more difficult when using discreet categories (e.g. how visually different are ‘A’ from ‘B’ compared to ‘Z’). Researchers found that the number of stimuli participants needed to maintain in VSTM decreased the precision of their answers (e.g. Adam, Vogel, & Awh, 2017; Olsson & Poom, 2005; Wilken & Ma, 2004). As a result, researchers have proposed that there is a trade-off between the number of items in VSTM and the resolution of the items.

2.2. THEORIES ON THE MODULATION OF ATTENTION

Several researchers (Bays & Husain, 2008; Donkin, Kary, Tahir, & Taylor, 2016; Frick, 1988; Wilken & Ma, 2004) have explained the trade-off between items in VSTM and precision through the resource model. They argue that memory capacity should be understood not in terms of a limited number of slots, but rather as a limited pool of cognitive resources that can be distributed across the visual field.

This would mean that there is a trade off in the number of visual representations and the mental resolution of the items in VSTM (see fig. 2 for visual representation). This could theoretically mean that it would be possible to divide the resources among an unlimited number of items, as long as the cognitive resources were able to divide into small enough amounts. Some resource models also argue that the cognitive resources can be divided unevenly between the items in VSTM (Bays & Husain, 2008), creating some items with a high mental resolution, and others with low. Donkin, Kary, Tahir, and Taylor (2016), showed that it was possible to manipulate the participants into using either the slot or the resources model. They did this by using either a fixed or varying number of stimuli in the display. They found that participants would use slot-like encoding if they could not predict how many stimuli were shown, while only about 50% used an optimal resource-like encoding when the number of stimuli shown were fixed. This indicates that participants prefer slot-like encoding even when it is suboptimal, but are still able to use resource-like encoding in the right circumstances.

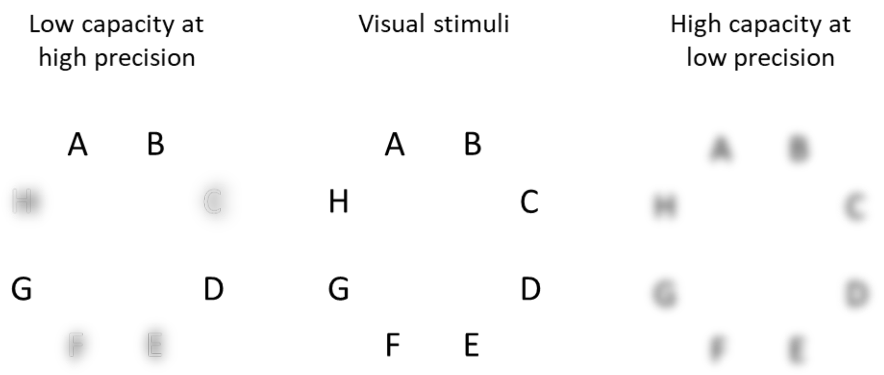

Fig. 2 Cartoon of the tradeoff between capacity and precision. The middle array shows the possible stimuli. The figure to the left show the letters A, B, D, and G are all stored as high precision items in VSTM whereas there are no firm representation of the other letters. The figure to the right show all the letters stored though only as low precision items.

Studies that provide empirical evidence for the resource model have, however, mainly used partial report experiments with simple stimuli (e.g. Bays, Catalao, & Husain, 2009; Bays & Husain, 2008; Donkin, Kary, Tahir, & Taylor, 2016; Donkin, Tran & Nosofsky, 2014; Whilken & Ma, 2004). Participants in partial report experiments are shown stimuli and then asked to report a subset of the stimuli or report whether a secondary stimuli were in the display. Determining whether a stimuli were in a display requires a comparison of the item in memory to one on the screen. This comparison might bias the participants to maintain all stimuli holistically in fewer slots. This change in memory strategy might be the reason why some participants use resource-like encoding in some experiments (Donkin, Kary, Tahir, & Taylor, 2016). I am later in my thesis therefore going to examine whether using other designs, like whole report, would find different results as whole report would require the participants to maintain each stimulus separately.

This difficulty in showing a resource-like encoding has resulted in several researchers arguing that there is a limited number of slots (e.g. Scolari, Vogel, & Awh, 2008; Zhang & Luck, 2008), with several differing opinions on the way participants are able to conduct a slot-resolution trade-off.

Zhang and Luck (2008) proposed that the visual memory capacity was pseudo fixed. They argued that while there was a seemingly hard limit on the number of items it is possible to hold in VSTM, several slots could work together to improve the mental resolution of stimuli. This collaboration between slots reduced the number of remembered items but increased the precision. The relationship between the number of items maintained in VSTM and the mental resolution of stimuli was tested in a subsequent study (Xie & Zhang, 2017a). The study was conducted using familiarity with Pokémon characters to modulate the VSTM capacity, while also employing a confidence scale as a measure of the mental resolution of the participant. They found that participants familiar with Pokémon had a higher VSTM capacity for Pokémon, but were unable to find any significant familiarity differences in their mental resolution (participant’s confidence in their answers). This was used to argue that familiarity (at least in this case) improved the VSTM capacity while showing no changes in the mental resolution of the stimuli (Xie & Zhang, 2017a).

Scolari, Vogel and Awh (2008) suggested that while participants might have a fixed number of slots in VSTM, the precise distribution of cognitive resources can be divided in whatever way is optimal (Lorence, Pratte, Angeloni, & Tong, 2014; Scolari, Vogel & Awh, 2008). Zhang and Luck (2008) used the apt metaphor (despite arguing against the model) of having a set number of cups (slots) and being able to divide juice (cognitive resources) between the cups in any way. Several researchers have supported this model by finding that participants are able to report some items more precisely than others (e.g. Adam, Vogel, & Awh, 2017; McElree, 2001; Olsson & Poom, 2005). Adam, Vogel, and Awh (2017) used whole report to find that memory precision fell with each additional stimulus in the display over the VSTM capacity, and that the first item had the highest precision, with each successive item getting more imprecise.

Bays, Catalao, and Husain (2009) have, however, pointed out that a slot model, where one or more slots can work together to increase precision, is fundamentally similar to the resource model. They further argue that there is a problem in partial report experiments, which impacts findings of several studies that previously found an upper limit to the VSTM capacity (e.g. Zhang & Luck, 2008; 2009). The problem is that participants need to not only remember the items shown, but also the probed location. They therefore argue that some of the ‘random guesses’ in previous experiments might instead be misremembering of locations.

This discovery of decreasing precision for each subsequent item is not new (McElree, 2001), and creates part of the basis for the hypothesis that there is only one memory slot in VSTM that can hold anything of precision (McElree, 2001; Olsson & Poom, 2005). Olsson and Poom (2005) found that participants were unable to remember more than one item in an experiment consisting of a coloured ellipse with another ellipse inside (see fig. 3). This ellipse varied along three continuous axes (shape of the big ellipse, shape of the small ellipse, and colour). The point of the argument made by Olsson and Poom (2005) was that the participants’ inability to use LTM was what resulted in the ‘true’ VSTM capacity. However, this low memory capacity might be explained by the results from Jackson, Linden, Roberts, Krigeskorte and Haenschel (2015), as they found that it was the discriminability, rather than visual complexity, that decreased the VSTM capacity. Jackson et al. (2015) also used figures where participants had little chance of using verbal strategies and found a VSTM capacity for high complexity stimuli with high discriminability.

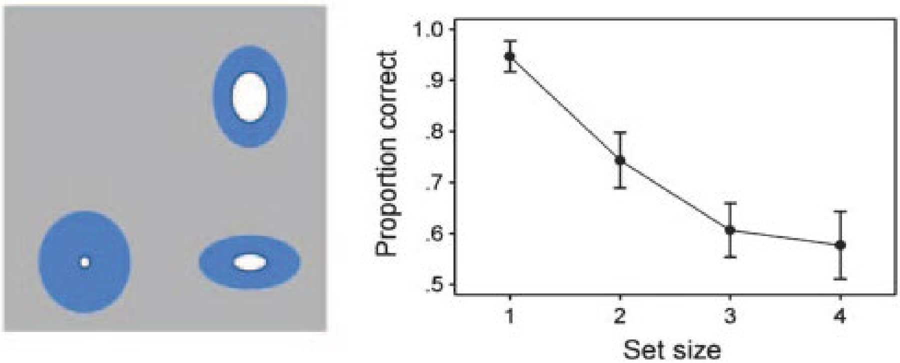

Fig. 3 Example of a stimuli display and graph showing performance (error bars denote 95% CI) from Olsson and Poom (2005)

The primary use of one experimental paradigm for model creation can create some problems (Donkin, Kary, Tahir & Taylor, 2016), such as the debate between the slot and resource models (e.g. Scolari, Vogel, & Awh, 2008; Wilken & Ma, 2004; Zhang & Luck, 2008). This debates whether the VSTM capacity should be understood as a number of slots capable of holding one item, or as a pool of recourses that can be divided across the visual field, allowing each item to be remembered at different resolutions. A good example of how the design might influence the strategy comes from Scolari, Vogel, and Awh (2008). In their study, participants are shown a visual display of several different stimuli categories (e.g. faces, geometric figures, etc.) for 1000 ms, followed by a change to the display after a 1000 ms blank display. Here, the optimal strategy would be to create a shallow representation of the entire display or to encode the stimuli into categories. This strategy of shallow representations seems to be the one that participants used in Scolari, Vogel and Awh (2008), as participants remembered an average of between 3.5-4.1 stimuli when the category changed, while between 0.5-1.6 stimuli when it did not change, depending on the condition. The within category change is far lower than in other studies, as a K of 0.5 for inverted faces is far below the memory capacity found in other studies (e.g. about 1.7 in Curby & Gauthier, 2007 despite using auditory suppression). Conversely, the optimal strategy would be to make a precise representation of all the stimuli in studies where participants must report what they saw/what was at a specific position.

2.2.1. TVA

An alternative theory to the preceding ones is the Theory of Visual Attention (TVA) (Bundesen, 1990). The TVA is a mathematical model developed by Bundesen (1990) and has been used in both basic (e.g. Dall, et al., 2016; Kraft, Dyrholm, Bundesen, Kyllingsbæk, Kathmann, & Brandt, 2013; Sørensen, Vangkilde & Bundesen, 2014; Ásgeirsson, Nordfang, & Sørensen, 2015) and clinical research (e.g. Caspersen, Petersen, Vangkilde, Plessen, & Habekost, 2017; Gögler, Willacker, Funk, Strube, Langgartner, Napiórkowski, Hasan & Finke, 2017; Habekost, 2015; Wiegand, Töllner, Dyrholm, Müller, Bundesen, & Finke, 2014).

TVA is based on the 'fixed capacity race model' (Bundesen, 1987; Shibuya & Bundesen, 1988). This race (and thereby TVA) model is a stochastic race between all available stimuli towards representation in VSTM (Bundesen, 1987). The race happens in two phases (see fig. 4). In the first phase all stimuli in the visual field are given a specific speed at which they are processed if they receive cognitive resources (Bundesen, 1990). This functions as a Poisson count model for each stimulus, where the likelihood that a stimulus is processed approaches 1 as time (t) increases (Bundesen, 1987; Kyllingsbæk, Markussen, & Bundesen, 2012). The second phase is the race itself, where cognitive resources are divided amongst the stimuli (Bundesen & Habekost, 2008).

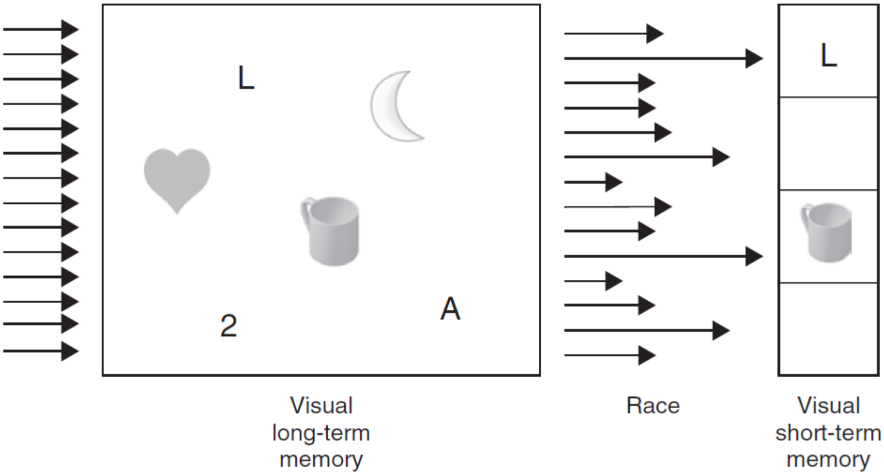

Fig. 4 Cartoon of the race model. The sensory evidence activates the long-term memory which in turn starts a race between the activated items. The VSTM has only a limited capacity and it is therefore only the fastest four items that are remembered (Bundesen & Habekost, 2008).

Each stimulus will have a Poisson distribution that can be written as:

$$p_n = \frac{v(i,j)^n (t - t_0)^n}{n!} e^{-v(i,j)(t-t_0)}$$

The probability that a stimulus receives a tentative categorization after n (number of tentative categorizations that stimulus i belongs to category j) counts (p_n), is the Poisson distribution's λ ($v(i,j)(t-t_0)$) divided by the factorial of n times e to the power of negative λ. This means that λ is the rate of processing that stimuli, i belongs to category j ($v(i,j)$) to the power of n times the time stimuli have been shown (t) minus the threshold for minimal perception (t_0) to the power of n (Kyllingsbæk, Markussen, & Bundesen, 2012).

This race between stimuli continues until either each slot of a limited number of slots in the STM is full, the stimuli are removed (Bundesen, 1987), or until the beginning of a new saccade (Bundesen & Habekost, 2008). While it is described as speed, it would be more correct to call it the chance of the stimuli being processed (Kyllingsbæk, Markussen, & Bundesen, 2012). The speed is therefore closer to bookies odds in a horse race rather than how fast the horse finishes the race. For the sake of simplicity, it is furthermore assumed that processing starts at the same time for all stimuli (t = 0), although it probably starts with each saccade (Bundesen & Habekost, 2008) (see fig. 5 for visual representation).

The probability that a specific stimulus will be chosen (P(i,j)) can be written as:

$$P(i,j) = P_1(i,j) + P_2(i,j) + P_3(i,j)$$

This probability is, how likely a specific stimulus is the one with the most counts ($P_1(i, j)$), plus the possibility that it is chosen while at capacity limit along with another stimulus and sharing the highest number of counts with only one slot left ($P_2(i, j)$), plus the probability that all counts are zero and the participant guesses at random ($P_3(i, j)$) (Shibuya & Bundesen, 1990).

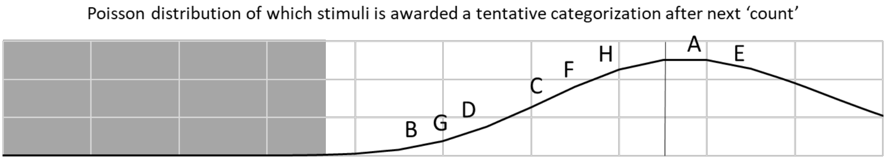

Fig. 5. Shows a visual representation of the Poisson distribution of receiving an extra tentative categorization. Everything in the greyed outside is before t_0, and there is therefore 0% chance that a perceptual unit have been encoded into *K* at that timepoint. There is still a chance that a perceptual unit is encoded into *K* at the timepoint next to t_0 ($t_0 + 1$), as it is probability calculations. Everything on the right side of the black line is already encoded but might still receive an extra tentative categorization after a 'count'.

The likelihood of a specific stimulus being processed is given by the Poisson intensity called the rate function (or the Poisson intensity), which is created for each stimulus in the visual field (Christensen et al., 2018). The rate function thereby acts as an indicator for how long it takes the stimuli to be stored in VSTM (Bundesen, 1990). This is done by calculating the likelihood of a specific perceptual unit x being processed as belonging to category i, which can be written as $v(x, i)$. It should be noted that perceptual categories are fairly broad and can be specific colours, shapes, and locations (e.g. the likelihood that 日 belongs to the category Chinese characters). This can be calculated as:

$$v(x,i) = \eta(x,i)\beta_i w_x / \sum_{z \in S} w_z$$

$\eta(x, i)$ is the sensory evidence that perceptual unit x is part of category i (Bundesen, 1990). Participants conduct a mental compresence between a mental template (category i) and the perceptual unit (perceptual unit x). The η is therefore highly influenced both by the visibility of the perceptual unit (Christensen, Markussen, Bundesen, & Kyllingsbæk, 2018), as well as the similarity between the mental template and the perceptual unit (Bundesen, 1990). This can be high, if e.g. x is presented as white on black or low, if e.g. x is presented as grey on light grey or there is something obscuring x. Template matching occurs when a perceptual unit is categorised as a specific "object". An example of this might be the character 'へ', which might be categorised as several different things. It might be categorised as a tilted 'L' if the person does not know Japanese, or it might be categorised as the Japanese hiragana 'he' if the person is familiar with Japanese. There would be no η

for 'へ' being Japanese if the person does not know Japanese hiragana and therefore does not have any mental template for the Japanese 'he'. A person with the same level of Japanese and Latin expertise would have a η for both Japanese and Latin, but the one for Japanese would likely be higher, as a study by Pelli, Burns, Farell, and Moore-Page (2006) suggests that the closer a stimulus is to its mental representation, the faster it is processed.

β_i is the perceptual bias that a perceptual unit is processed as category i, e.g. looking for a fire extinguisher in a burning house. Bundesen, Vangkilde and Habekost (2015) argued in a later paper that the β is made up of several features.

$$\beta_i = A * p_i * u_i$$

where A is alertness, p_i is the subjective prior probability that a perceptual unit will be part of category i (i.e. even native Japanese speakers are unlikely to report seeing 'へ' as 'he' if the task is to report tilted Latin letters), and u_i is the importance of identifying that specific feature. The influence of alertness does, however, differ between K and C, where a too high arousal A decreases K, whereas C increases even at high arousal (Bundesen, Vangkilde & Habekost, 2015; Sørensen, Vangkilde & Bundesen, 2014).

w is the attentional weight for a specific perceptual unit (w_i is the attentional weight for i), which is how likely it is that a specific perceptual unit is processed. Several researchers (e.g. Intriligator & Cavanagh, 2001; Kraft, Dyrholm, Bundesen, Kyllingsbæk, Kathmann, & Brandt, 2013) have argued that there is a processing advantage for stimuli in the upper hemisphere, as well as an improved memory capacity for stimuli presented across both left and right hemispheres.

This means that the speed at which a perceptual unit is processed (rate equation) can be given by the sensory evidence that a perceptual unit is a specific category (η), times the perceptual bias for a specific category (β_i), and then divided by the attention that specific perceptual unit gets (w). All this is divided by the sum of the attention all the other perceptual units gets ($\sum_{z \in S} w_z$). Is it therefore possible to use the rate function to create a hazard function for the likelihood that the perceptual unit x is processed as category i at time t:

$$p_E | K \ \geq n(S), t > t_0 = 1 - \exp\left[-v(x,i)(t - t_0)\right]$$

This is therefore the probability that element x is encoded into VSTM at time t, as long as t is longer than t_0 ($t - t_0$) and K is bigger or equal to the number of perceptual unit (n(S)). This probability can be written as 1 minus the exponential function for the hazard function of $v(x, i, t)$ ($-v(x, i)(t - t_0)$). While a hazard function is exponential, previous research has shown a better fit for TVA with a sigmoid fit (Dyrholm, Kyllingsbæk, Espeseth, & Bundesen, 2011; Petersen & Andersen, 2012). This was solved when it was proposed that t_0 varies from trial-to-trial (Dyrholm, Kyllingsbæk, Espeseth, & Bundesen, 2011).

The rate equation can therefore be divided into the bottom-up with the sensory evidence (η) for a specific perceptual unit, and the top-down influences on processing with the perceptual bias (β) and the pertinence value (π) as the current importance of a specific perceptual unit (Bundesen & Habekost, 2008). The top-down influence of the pertinence (π) is the bias of attending to perceptual units from a specific category, and is important for the amount of attention given to a specific perceptual unit (Bundesen, 1990). Nordfang, Dyrholm, and Bundesen (2013) quite effectively showed that neither contrast or singletons (Theeuwes, 1991) influence π, but π is rather influenced by the strength of a local feature (κ). Nordfang, Dyrholm, & Bundesen (2013) showed that the saliency of κ drew attention (e.g. a red singleton among blue L's) regardless of whether the participants were looking for a singleton or not. κ can therefore be seen as a non-negative general multiplier of attention, drawn towards a perceptual unit. Furthermore, κ gained more influence the more stimuli there were in the visual field, indicating that it influences the attentional weight of a perceptual unit (w) (Nordfang, Dyrholm, & Bundesen, 2013). κ and π can therefore be called the bottom-up and top-down bias for local features, where the pertinence is considered as a multiplier of η (Bundesen, 1990) rather than the overall salience effect of κ (Nordfang, Dyrholm, & Bundesen, 2013). This can be calculated through the weight equation:

$$w_x = \kappa_x \sum_{j \in R} \eta\,(x, j)\pi_j$$

The attentional weight of x is the strength of the local feature for x (κ_x) times the sum of all visual categories $\sum_{j \in R}$ for the sensory strength that perceptual unit x belongs to category j ($\eta(x, j)$) times the pertinence for j (π_j). K has traditionally been calculated as one less than the maximum number of items correctly remembered plus the probability that the participant remembered the maximum number of items (Bundesen, 1990). Dyrholm and colleagues (2011) recalculated K to include a larger trial-by-trial variance, where K was defined as the mean score, through this function:

$$E[K|\mathbf{m}] = \sum_{j=0}^{\infty} m_j \times j$$

It is of primary importance to define the normalized histogram for K as **m** and the probability that $K = j$ on a given trial as m_j. $E[K|\mathbf{m}]$ is the function for the estimated K on any given trial (**m**). This means that the estimated K on a given trial is the sum of the probability for each possible K. K is not the maximum number of perceptual units in VSTM (K_{max}), but rather the expected value of discrete random variables. K_{max} is the upper bound of **m** and cannot be less than the maximum number of trials reported on any given trial if it is assumed that the participants did not guess (Dyrholm et al., 2011). The probability that K on any given trial is lower than $K_{\max}$ ($K < K_{\max}$), is:

$$p_{inc}\frac{E[K] - K_{min}}{K_{max} - K_{min}}$$

This means that the independent probability for upper and lower bound is the estimated mean K minus $K_{\min}$ divided by $K_{\max}$ minus $K_{\min}$. There is therefore a 1 – p_{inc} probability that $K_{\max} - K_{\min} + 1$ number of tasks unrelated perceptual units is encoded into VSTM.

2.2.1.1 NTVA

NTVA, inspired by Hebb (1949), is the theory that visual stimuli are seen as the activation of several populations of neurons, which represent the encoding of a stimulus (Bundesen, Habekost, & Kyllingsbæk, 2005). Encoded stimuli are thereafter maintained in reverberating feedback loops from early processing through thalamus to prefrontal cortex (Usher & Cohen, 1999). An object is estimated to only be in VSTM if the topographical representation of that object is activated.

A neuron (or clusters of neurons) acts in a similar way to the neurons studied by Hebb (1949), and will only respond to one feature, termed the feature-i neuron. While the neuron might only respond to one feature, that feature might be the same features in different objects. It should be noted that feature i might be a gradient of a category, such as a shade of colour, where something can have different degrees of feature i.

Cells function in accordance with winner-take-all clusters, proposed by Grossberg (1976, 1980). The proposal in NTVA is that when a cluster of cells (or an object in VSTM) reaches an activation rate, they will inhibit other cells. This rate is dependent on the rate equation from TVA, where $\eta(x, i)$ is the total number of activated neurons when there is only one perceptual unit, while β_i is the activation rate (Bundesen, Habekost, & Kyllingsbæk, 2005). It should be noted however that both β and π are estimated to be late-stage processing, such as higher order areas in frontal, parietal cortex, and limbic system (Bundesen, Habekost, & Kyllingsbæk, 2005).

There are a number of processing bottlenecks or gates, which will only allow information through from a specific number of cluster of cells. Bundesen and colleagues (2005) suggest that the first of these gates is between V2/V3 and V4 as well as a further gate between V4 and the inferotemporal cortex.

Later studies (e.g. Wiegand, Töllner, Dyrhold, et al., 2014; Wiegand, Töllner, Habekost et al., 2014) have found associations between K and C, together with EEG patterns that have already been studied thoroughly. *K* seems to correlate with the well studied EEG phenomena called Contralateral Delayed Activity (CDA; Vogel & Machizawa, 2004). CDA is the larger negative activation in the hemisphere contralateral to where a stimulus is shown, compared to the ipsilateral. This difference happens around 200ms after stimuli presentation, and occurs mostly in the posterior parietal and lateral occipital regions. The strength of this negative difference is indicative of the number of items in VSTM. This difference seems to indicate the number of items in VSTM regardless of stimulus specific expertise (Xie & Zhang, 2018) or memory load (e.g. Gao, Yin, Xu, Shui, & Shen, 2011).

The neural correlate of *C* does not correspond only to one EEG parameter, but rather to several, such as N1 (Wiegand, Töllner, Dyrhold, et al., 2014; Wiegand, Töllner, Habekost et al., 2014) and alpha frequency (Hilla, von Mankowski, Föcker, & Sauseng, 2020), and this is also reported as a measure for processing speed in previous literature (Klimesch, 1999; Klimesch, Doppelmayr, Schimke, & Pachinger, 1996; Klimesch et al., 2004).

2.3. TVA AS METHOD

The mathematical nature of TVA makes it possible to plot memory capacity at different *t*. This will, in most cases, create a sigmoid function for the number of encoded stimuli (Dyrholm et al., 2011). This function allows for the disentanglement of several cognitive functions, including those central to this thesis: the threshold for perception (t_0); processing speed (*C*), and VSTM capacity (*K*) (Bundesen, 1990). As previously noted, *C* dictates the speed by which a perceptual unit is encoded into an empty VSTM (Bundesen, Vangkilde, & Petersen, 2015), and can therefore be calculated as:

$$C = \sum_{x \in S} v(x) = \sum_{x \in S} \sum_{i \in R} v(x, i)$$

This means that the processing speed is the sum of the rate equation ($v(x, i)$) for all categories ($i \in$ R) across all perceptual units ($x \in$ S) (Bundesen, Vangkilde, & Petersen, 2015). This can also be described as the slope of the sigmoid function at t_0. The number of items encoded into VSTM will slowly reach an asymptote as *t* increases. This asymptote has been shown to be a good measure for the VSTM capacity (*K*) (see fig. 6 for TVA fit example) (Bundesen & Habekost, 2008). It should, however, be noted that increasing the exposure rate over a specific point increases the likelihood of participants being able to chunk information (Sperling, 1967).

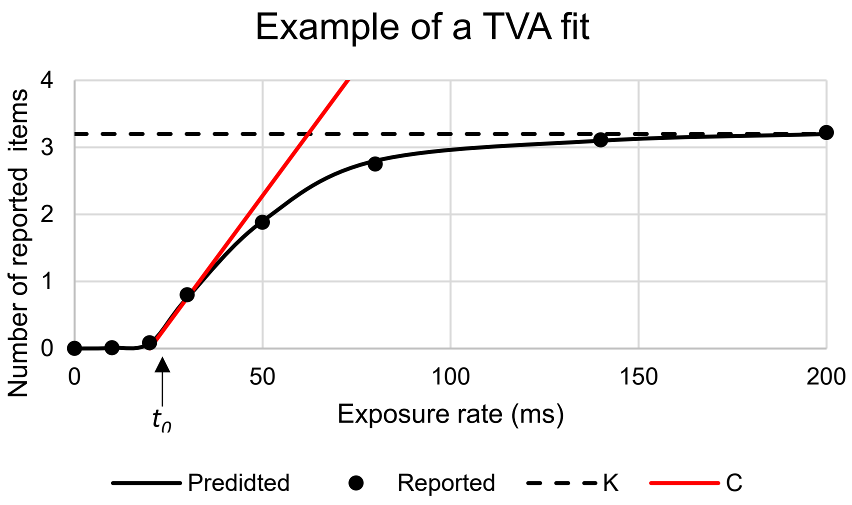

Fig. 6. Example of a TVA fit (data from Paper 2). The black dots represent the mean correctly reported stimuli at the different exposure rates, while the black line is the TVA model fitted to the data. The arrow points to the threshold of minimal perception (t_0), while the red line represents the processing speed (C). The black dotted line following the asymptote of the curve indicates the upper limit of the VSTM capacity (K)

Several studies have been conducted to validate K, C, and t_0, and have found that all three have a high reliability (Bundesen & Habekost, 2008). This disentanglement of cognitive functions has given more nuance to the cognitive profile of psychiatric disorders such as ADHD (Caspersen, Petersen, Vangkilde, Plessen, & Habekost, 2017), psychostimulants (Finke et al. 2010) and depression (Gögler, Willacker, Funk, Strube, Langgartner, Napiórkowski, Hasan & Finke, 2017).

Here the focus will be directed towards K, C, and t_0, as these are the focus in the three papers which create the basis for this thesis.

One of the problems with the TVA equations is that they are not comprehensive in explaining all the data. K is influenced by stimulus specific expertise, which is not very problematic as there is no equation for K, but this in itself is a problem. The problems start in earnest when you consider C and t_0. C influences K, but only to a degree, as psychoactive improvements to C in a low C population increases K, but not in an average-high C population (Finke, Dodds, Bublak, Regenthal, Baumann, Manly, & Müller, 2010). This becomes even more problematic when stimuli contrast is decreased, as this decreases t_0 but not K or C. Another problem is that if t_0 is the fastest v_x this would suggest that the faster C is, the lower t_0 will be, but the complexity experiment and the Japanese experiment 2 both influence C while not influencing t_0. These problems do not correspond to the TVA equations, as there are no variable influences on t_0 and not C.

The independence of *C* and *K* is well argued to be a difference in early and late processing. *C* seems to have a strong connection to early processing and occipital N1 amplitude. Meanwhile, *K* has been shown to have a connection to the contralateral delay activity (Wiegand, et al., 2012). The relationship between *K* and CDA seems to support its validity as a measure of VSTM, as several authors have argued that CDA is one of the best neural markers for VSTM (see e.g. Luria, Balaban, Awh, & Vogel, 2016 for review).

A study using TVA has shown that it is possible to use methylphenidate and modafinil to improve *C* for participants with a *C* lower than the group median. There was a further significant *K* influence of modafinil for participants with a *K* lower than the group median. Additionally there was a correlation between attention and *C* (Finke et al., 2010). Nicotine seems to reduce t_0 and *C* causing the start of stimuli processing to happen earlier, but at a slightly slower speed, while showing no influence on *K* (Vangkilde et al., 2011).

C improvements might be explained by proposed changes to TVA β made by Bundesen, Vangkilde and Habekost (2015). They proposed that β is made up of $A*p_i*u_i$, where A is alertness, p_i is the subjective prior probability that object *i* is part of category *p*, while u_i is the importance of identifying feature *i.* The p_i might therefore be the stimuli specific expertise that influences the speed of the *C*.

C is commonly estimated in items pr. second (Hz). The actual processing speed at a given exposure rate would therefore be: $C/(\frac{1000\,ms}{exposure\ rate\ in\ ms})$. A problem arises when using items rather than perceptual categories pr. second. There are several indications that VSTM is influenced by expertise (e.g., Dall et al., 2016; Sørensen & Kyllingsbæk, 2012), which suggests that this is partly due to the number of slots or resources an object takes. Furthermore, if *C* is the number of processed objects per second, then *C* would have to decrease drastically to have an influence on *K,* as *C* is shown in several studies to be over 10 times as high as *K* (e.g. Finke, et al., 2010; Vangkilde, et al., 2011; Wiegand, et al., 2012; Dall et al., 2021).

2.4. METHODOLOGICAL CONSIDERATIONS FOR TVA

While TVA has several clear benefits, there are some methodological considerations that need to be addressed in TVA.

Measuring t_0 becomes important as *C* being the slope in t_0 (Matthias et al., 2010). This can make the slope difficult to estimate correctly if its starting point is inconsistent. The problem arises when the spaces between the exposure rates around t_0 are too large. This can be especially problematic if the stimuli are shown on a 60Hz monitor, as a t_0 below 16ms is not uncommon in high contrast stimuli. It is possible to lock t_0 at a specific point (often at 0ms) if t_0 is negative (Kyllingsbæk, 2006), or to change the stimuli durations to match the participant after some pre-test trials (Gögler, Willacker, Funk, Strube, Langgartner, Napiórkowski, Hasan & Finke, 2017).

Another consideration is problems arising from the mask being too weak. A weak mask can result in participants seeing through the mask and perceiving stimuli at the shortest exposure rates (usually below the theoretical t_0). This can create situations where t_0 is calculated to be below 0ms (before stimuli is shown), and while some branches of psychology might hail seeing stimuli before it is shown as proof of the ability to predict the future, it seems theoretically unlikely.

2.4.1. DATA MODELLING

Data is traditionally modelled either by hand (e.g. Bundesen, 1990) or through the LibTVA (e.g. Dyrholm, Kyllingsbæk, Espeseth, & Bundesen, 2011) for MatLab, or via a custom mathematical model (e.g. Christensen, et al., 2018). LibTVA is the only data model that has been standardized and is used in the standardized TVA paradigm CombiTVA (Vangkilde, Bundesen, & Coull, 2011). LibTVA aims to create a model for attention, which can then be used to calculate the different parameters (e.g. K, C, t_0, etc.). Previous studies have tried comparing models (e.g. Christensen, et al., 2018; Dyrholm, Kyllingsbæk, Espeseth, & Bundesen, 2011; Sørensen, Vangkilde, & Bundesen, 2015), and all have found that adding further parameters (e.g. model shape, or trial-by-trial variance) improves the model. It is the exception rather than the rule that studies use resources on testing the model.

It is common practice to add variables to TVA models to better fit them to the data. These new TVA models are problematic however in relation to the number of variables in the model compared to the number of data points (exposure rates). Having more variables than data points drastically increases the risk of overfitting the TVA model. It is important in these circumstances to remember the paraphrasing of George Box (1976; 1979): “All models are wrong, but some are informative”. There are steps researchers can take to correct for the possibility of overfitting when comparing models by using methods such as LASSO (Santosa & Symes, 1986; Tibshirani, 1996) and ridge regression (Phillips, 1962; Tikhonov & Arsenin, 1977), although both of these steps are not implemented in LibTVA and require training data. An alternative method is to examine whether additional variables add or subtract from the model’s explanatory value. This is possible by using either the Akaike information criterion (AIC) (Akaike, 1973) or the Bayesian information criterion (BIC) (Schwarz, 1978). The different information criterion is calculated by comparing the log likelihoods of different models. This requires comparing several TVA models with a different number of parameters and choosing the one with the most explanatory value. A final method is by implementing a χ^2 goodness of fit test between the observed and their correspondent predicted points in the TVA model.

2.4.2. TVA EXPERIMENTAL DESIGN

TVA has traditionally been used with whole report, partial report (Sperling, 1960), and change detection studies (Pashler, 1988; Phillips, 1974). These different designs can identify specific cognitive parameters. While whole and partial report experiments were used in the work of early psychology researchers (e.g. Cattell, 1885; Wundt, 1899), they gained popularity after a series of studies by Sperling in the 1960s (e.g. Sperling 1960; 1963; 1967). Participants in both types of paradigms are presented with various stimuli that they must remember and report. While the number of items are often above the memory capacity of the participant (e.g. Kraft, et al., 2013; Sperling, 1960), some authors have used fewer stimuli to examine the precision and processing speed of a stimulus (e.g. Adam, Vogel, & Awh, 2017).

The paradigms differ on what the participants are requested to do with the remembered stimuli. The participants in a whole report paradigm are required to report all the stimuli they have remembered, whereas participants in partial report experiments must report a subset of the shown stimuli (Sperling, 1960). The original studies by Sperling (1960) used sound as the indicator of which line, out of three, participants had to report. In TVA this has commonly been replaced with the colour of stimuli (e.g. Duncan, Bundesen, Olson, Humphreys, Chavda, & Shibuya, 1999; Vangkilde et al., 2011). This is done by showing stimuli in more than one colour, where only one of the colours should be remembered, while the rest of the stimuli serve as distractors. In both paradigms, stimuli are often presented within one saccade (e.g. Adam, Vogel, & Awh, 2017; Dall, et al, 2016; Dyrholm, Kyllingsbæk, Espeseth, & Bundesen, 2011). However, some researchers have found good results with longer durations in participants with slowed processing (e.g. Depression: Gögler, Willacker, Funk, Strube, Langgartner, Napiórkowski, Hasan & Finke, 2017), while short stimuli durations were originally used to examine immediate memory (Sperling, 1960). Later researchers have suggested that brief information overload might prevent chunking and therefore gives an estimate of VSTM capacity (e.g. Bundesen, 1990; Cowan, 2001). Examining VSTM in this way, makes it possible to calculate K, C, t_0, and v, whereas it is also possible to calculate the influence of distractors in partial report paradigms (Kyllingsbæk, 2006). As the basic element of showing stimuli and asking what was shown is fairly simple, some researchers have used several different stimuli arrays, such as horizontal/vertical (e.g. Habekost & Bundesen, 2003), grid formation (e.g. Sperling, 1960) or circular (e.g. Kraft, et al., 2013).

Masking is used in TVA studies when it is important to control the effective exposure rate instead of calculating stimulus decay (Kyllingsbæk, 2006). Pieron (1925) originally coined the term ‘masking’, although the process of masking has been used for far longer (e.g. Baxt, 1871 as cited in Baxt, 1982; Stigler, 1910; Wertheimer, 1912; Kahneman, 1968). Masking a stimulus is the process of presenting a different stimulus before (forward masking), during (paracontrast and metacontrast), and/or after stimulus (backwards masking) onset (Boynton & Kandel, 1957; Kahneman, 1968). While there are many types of masking (e.g. light: Crawford, 1947; Sperling, 1964; visual noise: Kinsbourne & Warrington, 1962; metacontrast: Stigler, 1910; structure: Breitmeyer & Ganz, 1976; visual illusion: Crouzet, Overgaard, & Busch, 2014; Di Lollo, Enns, & Rensick, 2000; Hirose & Osaka, 2009), this thesis will focus on structure/pattern masking. One of the principal goals of masking is to precisely control the stimulus duration, as there are afterimages both on the retina and the screen (Sligte, Scholte & Lamme, 2008). A stimulus presented right after another will interfere with the afterimage of the first (Averbach & Coriell, 1961). How well a mask interferes with and/or stops the processing of a stimulus refers to its strength.

The mask itself has an influence on how strong it is. This was shown in an experiment where randomly organised black and white squares were a more effective mask compared to ordered squares/dots (see fig. 7) (Coltheart & Arthur, 1972).

Fig. 7. Shows the difference in mask strength between ordered and randomly placed black and white dots (Coltheart & Arthur, 1972).

Blalock (2013) further argued that target mask similarity improved the strength of the mask when remembering coloured squares. The dissimilar mask consisted of a 4 X 4 matrix with an even number of black and white cells. The similar mask consisted of colours in a 2 X 2 checkerboard pattern. It could be argued that the author did not take stimuli salience into account when selecting similar/dissimilar stimuli, as luminance difference has a high salience (Krüger, Tünnermann, & Scharlau, 2017).

The mask strength is also influenced by the duration of the mask (Sperling, 1960b) as well as how fast it was presented after the stimuli disappeared (Baker, 1963; Kinsbourne & Warrington, 1962). Sperling (1960b) found that a mask presented for 500 ms was stronger than one presented for 50 ms. The strength of the mask is not just influenced by the duration of the mask, but also by the stimulus onset asynchrony (SOA) (the pause between when the stimuli is removed and when the mask appears), where the mask is strongest when SOA is 0 ms (Kinsbourne & Warrington, 1962).

2.4.3. STIMULI

TVA has, as previously noted, been divided into whole and partial (Sperling, 1960) report, and change detection (Pashler, 1988; Phillips, 1974) studies, with clear differences between the types of stimuli used in the different experimental designs. The main stimuli used in TVA whole and partial report experiments has been Latin letters, where participants should either name the stimuli or find them on the keyboard (Bundesen & Habekost, 2008; Kyllingsbæk, 2006). Change detection studies have consisted of both pictures (e.g. Sørensen & Kylligsbæk, 2012) and non-Latin graphemes (e.g. Sørensen & Kyllingsbæk, 2009). It is, however, relevant to give a short introduction to both Chinese characters and Japanese hiragana, as they are used in the papers this thesis is based upon.

The Chinese writing system is strictly speaking not a logographical language as characters can be combined. The smallest component of the characters - called radicals - are logographical (Woo, 2004). These radicals can be combined to make a new character, or be placed either next to or above another character to create a compound character (Crystal, 1987). The pronunciation of Chinese characters is usually logosyllabic and so for the most part only have one syllable. It is possible to combine several logosyllabic characters into a multisyllabic word (Woo, 2004). Chinese characters are created through a number of strokes with the pen. The official number of pen strokes are that character's stroke count. The stroke count of a character's main radical has traditionally been used to determine its place in the dictionary (Woo, 2004). Research indicates that radicals dictate the processing speed of the compound characters (Taft & Zhu, 1997).

Japanese hiragana is one of Japanese's several writing systems, and is comprised of 46 syllabic characters (Crystal, 1987). The other writing systems include another syllabic language in the form of katakana, a logographical writing system inspired by Chinese called kanji, the Latin alphabet, and Arab numbers (Crystal, 1987). Hiragana can be used to write any kanji, though in modern day it is most commonly used either in grammatical writing or as a way to spell uncommon words (Igarashi, 2007).

2.4.4. SUMMARY

I chose a whole report paradigm for all three experiments as it afforded studying the entire content of the participant's memory. TVA, in similar vein, made it possible to modulate distinct cognitive functions at the same time. Making it possible to pinpoint the precise influence of expertise on visual short-term memory and attention.

CHAPTER 3. PAPER SUMMARY

3.1. SUMMARY OF PAPER 1: CATEGORY SPECIFIC KNOWLEDGE MODULATES CAPACITY LIMITATIONS OF VISUAL SHORT-TERM MEMORY

This paper (Dall, Watanabe, and Sørensen, 2016) examines the influence of expertise on VSMT capacity using a whole report paradigm. Several researchers have demonstrated that stimulus specific expertise improves VSTM capacity (e.g. Curby, Glazek, & Gauthier, 2007; Gobet & Simon, 1996; Jackson & Raymond, 2005; McElree, 2001; Olsson & Poom, 2005; Scolari, Vogel & Awh, 2008; Sun, Zimmer, & Fu, 2011; Sørensen & Kyllingsbæk, 2014; Xie & Zhang, 2017a, 2017b; Zimmer, Popp, Reith, & Krick, 2012). Despite this overwhelming agreement, there are only a limited number of studies that report effects of stimulus specific training (Sørensen & Kyllingsbæk, 2014; Xie & Zhang, 2017a, 2017b), where most of the training occurred at an early age. Studies that have attempted training adult participants report no long-term improvement in memory capacity for the trained stimuli (Chen, Eng, & Jiang 2006; Olson & Jiang, 2004). Dall and colleagues (2016) suggest that the failure of studies in training participants to have long-term improvements in memory capacity might be caused by the degree of motivation to learn on the part of the participants. Bays and colleagues (2008) have further pointed out that most research within this field used partial report, thereby giving participants the additional cognitive task of not only remembering the stimulus but also remembering its location. This lack of replication in other designs was commented upon by Donkin and colleagues (2016), who argued that a design might condition participants to use a specific memory strategy.

As a result, paper 1 uses naturally trained participants to examine the influence of expertise on VSTM capacity, as well as extending results from previous studies from using partial report experiments to whole report experiments.

We attempted to examine this by utilising three conditions with stimuli at different levels of expertise: Simple line drawings (Pictures) from the Snodgrass and Vanderwart (1980) picture set, of which all participants had low expertise; letters from the Latin alphabet (Latin), of which all participants had a high level of expertise; and Japanese Hiragana (Hiragana) where the participants had one of three levels of expertise. The three levels of expertise for Hiragana were: Danish university students who had never studied Japanese or other Asian languages (Novices); Danish university students studying Japanese at the university (Trained); and Native Japanese university students in Japan (Experts). We ran a whole report experiment where the participants were shown six stimuli from one of the three conditions (Picture, Latin, and Japanese) at either 10 ms, 20 ms, 30 ms, 50 ms, 80 ms, or 200 ms, followed by a random-walk noise mask. The participants then used the mouse to select which stimuli they remembered from a grid with all the stimuli from that condition.

Data was modeled using TVA (Bundesen, 1990) using the LibTVA toolbox for MatLab (Dyrholm, Kyllingsbæk, Espeseth, & Bundesen, 2011). The random-walk noise mask we used was too weak, giving negative t_0 which forced us to lock t_0 at 0, thereby preventing us from modelling t_0 and C. We found that there was no significant difference between the groups in either the Picture $F(2,40) = 0.624$, $p = .541$ or the Latin condition $F(2,40) = 1.625$, $p = .21$. We did, however, find a significant difference between the groups in the Hiragana condition $F(2,40) = 9.049$, $p < .001$ (Novice < Trained < Experts), while we were unable to show any significant difference between the native language of the Novice and Trained (Latin) compared to the Experts (Hiragana) (see fig 8).

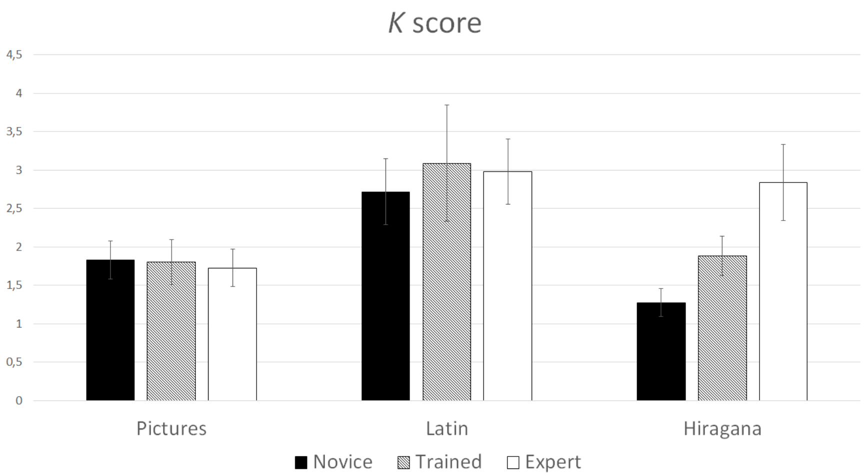

Fig. 8 Figure for the memory capacity (K) from Dall and colleagues (2016) of the three groups (Novices in black, Trained in grey, and Experts in white) in the three different conditions (Picture, Latin, and Hiragana). Error bars denote the standard deviations.

This large effect of Hiragana expertise supports previous research that suggests there is an effect of expertise on VSTM (e.g. Curby, Glazek, & Gauthie, 2007; Gobet & Simon, 1996; Jackson & Raymond, 2005; McElree, 2001; Olsson & Poom, 2005; Scolari, Vogel & Awh, 2008; Sun, Zimmer, & Fu, 2011; Sørensen & Kyllingsbæk, 2014; Xie & Zhang, 2017a, 2017b; Zimmer, Popp, Reith, & Krick, 2012). It also extends the results to whole report, supporting the suggestions of Bays and colleagues (2008) by way of circumventing the extra cognitive requirement to remember the positions of stimuli. The influence of expertise on K is hypothesized to be due to a strengthening of the LTM representations. This would mean that there is a matching between mental templates in LTM with visual sensory information. This is also supported by previous research that suggests that improvements to K are either due to an improvement in the mental resolution of stimuli (e.g. Scolari, Vogel, & Awh, 2008), or caused a reduction in the confusability between stimuli (Jackson, Linden, Roberts, Kriegeskorte, & Haenschel, 2015).

3.2. SUMMARY OF PAPER 2: VISUAL SHORT-TERM MEMORY AND ATTENTION: AN INVESTIGATION OF FAMILIARITY AND STROKE COUNT IN CHINESE CHARACTERS

Here we examined the influence of familiarity and stroke count for Chinese characters (Dall, Wang, Cai, Chan, and Sørensen, 2021). We did this by categorising the characters based on their frequency of use and stroke count. The study was inspired by the proposal of a relationship between visual complexity and the number of items that can be maintained in VSTM, with Chinese characters being more visually complex than letters (Alvarez & Cavanagh, 2004).

This understanding of relative complexity has been challenged by several other researchers (e.g., Curby et al., 2009; Dall et al., 2016; Jackson & Raymond, 2005; Sun et al., 2011; Sørensen & Kyllingsbæk, 2012; Xie & Zhang, 2017b). Jackson and Raymond (2005) proposed that visual complexity should be understood not only as physical complexity, but also as perceived complexity. Physical complexity would be measured in the number of meaningful lines, intersections, etc. (Jackson et al., 2015), whereas perceived complexity refers to how familiar the stimuli are to the observer. We tried to examine this relationship by categorizing Chinese logographical characters based on both familiarity and stroke count.

Chinese characters are ranked according to the number of strokes to determine physical complexity, while familiarity was estimated through a character frequency list, indicating how commonly the character is used. Most Chinese characters are made up of a combination of several characters. We therefore used Chinese radicals, which consist of 214 characters that are roughly equivalent to an alphabet. The individual radicals have, unlike Latin letters, a separate meaning (Woo, 2004).

We divided 60 of these Chinese Radicals into four groups, based on frequency of use and stroke count, resulting in 15 high stroke high frequency (HS-HF) radicals, 15 high stroke low frequency (HS-LF), 15 low stroke high frequency (LS-HF) radicals, and 15 low stroke low frequency (LS-LF) radicals. We defined the high and low stroke characters as having between 7-17 (Mean = 9.00, SD = 2.37) and 2-4 (Mean = 3.40, SD = 0.68) strokes respectively. The high frequency characters were defined as being amongst the 1000 most common characters (89.12% cumulative frequency, with a mean frequency of 489.47, SD =244.75). The low frequency characters were chosen from amongst the 2000th+ most common characters (cumulative frequency at 97.13% or above, with a mean frequency of 4602.00, SD = 2073.26) (Da, 2004), as functional literacy has been reported to fall within the 2000 most common words (Crystal, 1987). While the cumulative frequency sounds high, it includes characters for words like 殳 (meaning ‘lance’) or 鹿 (meaning ‘deer’).

Thirty-seven mainland Chinese participants were recruited from the Beijing Forestry University (thirty-three of which were female) and were tested in modern standard

Mandarin. We used university students to ensure that the participants had a uniformly high level of Chinese expertise.

The experiment was a whole report experiment. A trial stared with a white fixation dot at the centre of the screen for 500 ms. Stimuli from one of the four conditions was briefly presented [10 ms, 20 ms, 30 ms, 50 ms, 80 ms, 120 ms, or 200 ms] at 4 or 8 possible stimuli locations, followed by a random pattern mask (Vangkilde, Bundesen, & Coull, 2011; Sørensen, Vangkilde, & Bundesen, 2015). Participants then had to use the mouse to select the characters they saw from a 4 X 15 grid with stimuli from all four conditions (in the order of LS-HF, LS-LF, HS-HF, and HF-LS). No participant reported noticing that the stimuli in a trial were from only one of the four conditions. This was further supported by all participants reported seeing stimuli from more than one condition on a single trial indicating that the participants did not, even unconsciously, only select stimuli from a line where they already had selected one. The participants were shown a feedback screen at the end of each trail, indicating how many correct answers they reported. The entire experiment was comprised of 10 blocks of 140 trials (35 trial for each condition) as well as one block of practice trials at the beginning of the experiment. The duration of the entire experiment, including instructions and breaks, lasted approximately 3 hours.

Participants were additionally asked to only report stimuli when they were reasonably sure of what they had seen, and were informed that this equated to approximately 80 – 90 % certainty when making a decision. They were shown the percentage of correct responses at the end of each of the 10 blocks and were asked to maintain this value between 80 – 90 %. This was to homogenize the participants guess rates and prevent participants from reporting well under their K. An analysis showed that participants were able to keep within an 80 – 90 % correct rate. Data was modelled according to the Theory of Visual Attention (Bundesen, 1990), using the LibTVA library (Dyrholm, Kyllingsbæk, Espeseth, & Bundesen, 2011) for MatLab.

There was no significant t_0 difference between the conditions (see fig. 9). We were further unable to show any significant difference between the two high frequency conditions (HS-HF vs. LS-HF), as well as between the two low frequency condition (HS-LF vs. LS-LF) in K or C. There was both a significant K and a C frequency difference between the high frequency and low frequency conditions (see fig. 10 – 11).

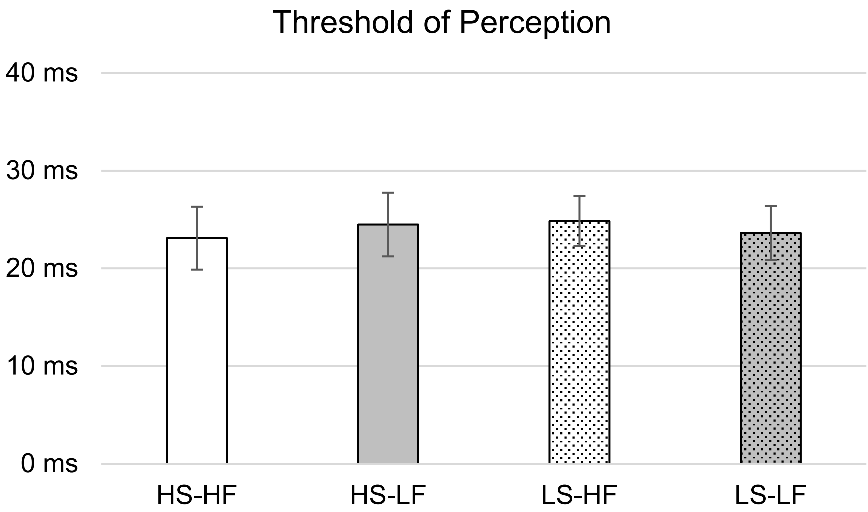

Fig. 9 Threshold for perception (t_0) between the conditions in Paper 2. Error bars denote a 95% confidence interval. HS-HF = High-Stroke High-Frequency condition (white), HS-LF = High-Stroke Low-Frequency condition (grey), LS-HF = Low-Stroke High-Frequency condition (white, hatched), LS-LF = Low-Stroke Low-Frequency condition (grey, hatched).

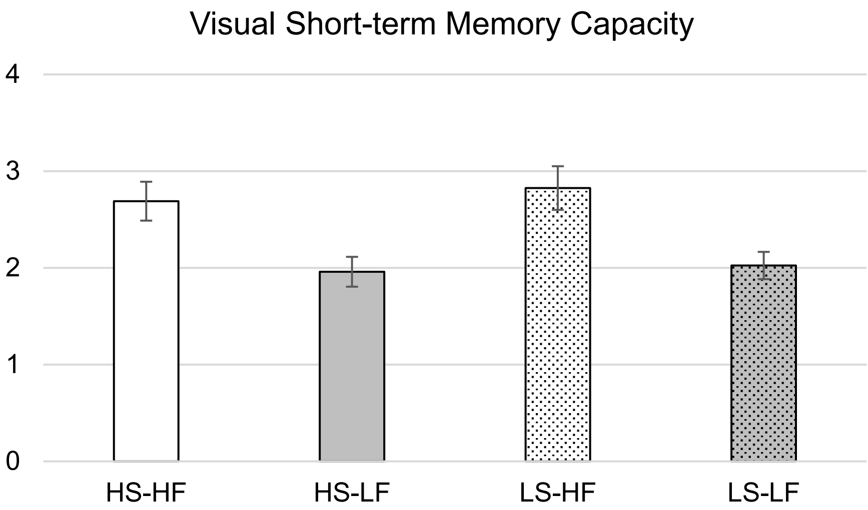

Fig. 10 The visual short-term memory capacity (K) between the conditions in Paper 2. Error bars denote 95% confidence interval. HS-HF = High-Stroke High-Frequency condition (white), HS-LF = High-Stroke

Low-Frequency condition (grey), LS-HF = Low-Stroke High-Frequency condition (white, hatched), LS-LF = Low-Stroke Low-Frequency condition (grey, hatched).

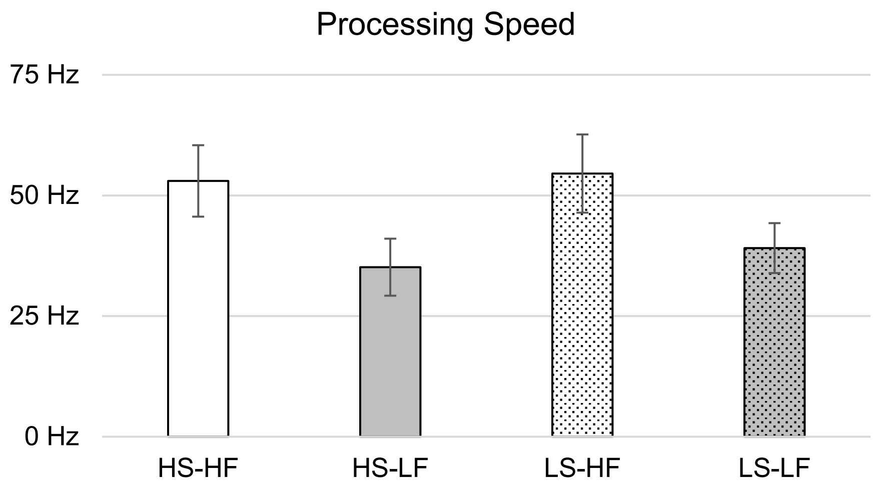

Fig. 11 The processing speed (*C*) between the conditions in Paper 2. Error bars denote 95 % confidence interval. HS-HF = High-Stroke High-Frequency condition (white), HS-LF = High-Stroke Low-Frequency condition (grey), LS-HF = Low-Stroke High-Frequency condition (white, hatched), LS-LF = Low-Stroke Low-Frequency condition (grey, hatched).

This suggests that frequency of use has a significant effect on both how fast and how many characters participants are able to remember, while the stroke count appears to have no significant influence on either the memory capacity or the processing speed. This supports previous research that reports expertise influences both in *K* (e.g. Curby et al., 2009; Dall et al., 2016; Jackson & Raymond, 2005; Sun et al., 2011; Sørensen & Kyllingsbæk, 2012) and *C* (Pelli, et al., 2006; Xie & Zhang, 2017b).

This is more likely to indicate that limited expertise with stimuli negates the influence of physical complexity, rather than that the physical complexity of stimuli had no effect. As Sun and colleagues (2011) demonstrate, stroke count negatively influencing the number of stimuli were remembered by native German participants with no expertise in Chinese.

This is supported by Jackson and colleagues (2015), who argued that expertise improves mental categories and therein reduces confusability. Swan, Collins, and Wyble (2016) similarly suggested that the resolution of a specific part of a stimulus is dependent on how ‘important’ the feature is. The naïve participants used by Sun and colleagues (2011) would, in this case, not have any mental category for the individual stimuli, leading to less optimal judgments on what are the important features, which might increase confusability between the stimuli.

3.3. SUMMARY OF PAPER 3: EXPERTISE INFLUENCE SHORT-TERM CAPACITY FOR COLOURS BUT NOT THE PRECISION OF REPRESENTATION

This paper (Dall, Chan, and Sørensen, *in prep*) studies how expertise within continuous categories influences the relationship between memory capacity and precision, using a whole report experiment, the purpose of which was to examine how expertise influences continuous categories. Numerous prior investigations have found that expertise improves both *K* (e.g. Curby et al., 2009; Dall et al., 2016; Dall et al., 2021; Jackson & Raymond, 2005; Sun et al., 2011; Sørensen & Kyllingsbæk, 2012) and *C* (e.g. Dall et al., 2021; Xie & Zhang, 2017b), where participants did not have time to encode the entire display. A few studies report that expertise with faces improved stimuli precision while not effecting *K* in experiments with a long encoding time (above 150 ms/item), (Lorenc et al., 2014; Scolari et al., 2008; Zhou et al., 2018). This resembles an observation made by Ye and colleagues (2019), who found that high preforming participants improved *K* when stimuli had a fast exposure rate (50 ms/item) but improved their precision during long exposure rates (100 ms/item).

We designed a whole report experiment examining the influence of expertise on continuous stimuli (colours) using high exposure rates. By doing this, we could combine a TVA design with a model for precision, which would provide additional information about processing speed and threshold for perception.

Stimuli were divided into discreet (Latin letters) and continuous stimuli (colours). The colours were selected from a CIE L^*a^*b from a 360° colour wheel centred at L = 76.65, a = 50, and b = 0, varying along a and b, which were divided into 20 colours with 18° between each colour. We modulated expertise by recruiting five participants from the School of Visual Arts from the Royal Danish Academy of Fine Arts as high expertise participants (Experts), and seven psychology students at Aalborg University, who reported not having any formal training in painting, graphical design or other visual arts (Novices).

Participants were briefly (10, 20, 30, 50, 80, or 200 ms) presented with either two, four, or six stimuli. This was followed by a rotating circle made up of the colours in the colour wheel, shown for 500ms. The participants then had to report as many stimuli as they could remember. They did this by selecting the position of a stimuli they could remember and either typing the letter, or using a colour wheel in the centre of the screen, depending on the condition (see fig. 12).

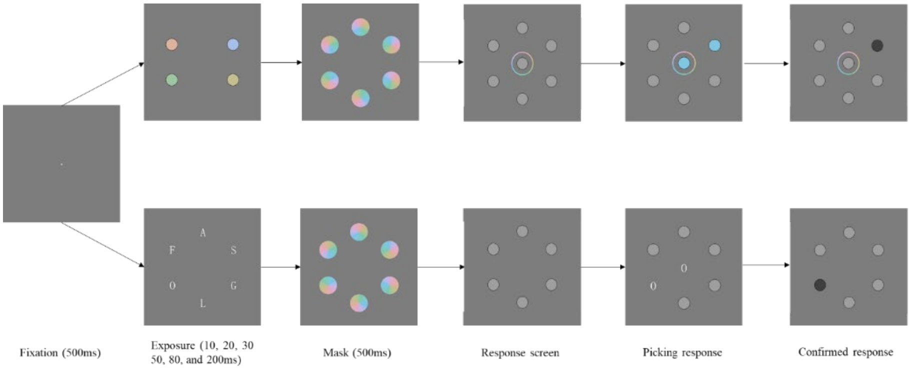

Fig. 12 An example of a colour (above) and letter (below) trial. The participants were shown either two, four, or six stimuli followed by a mask. The participants then have to use the mouse to select a position of a stimulus they could remember. The position turns dark grey after the participants have given their report.

Split-plot ANOVAs found that colour expertise did not influence the precision of reported stimuli $F(1,10) = 0.013$, $p = .825$; $\eta_p^2 = .001$ (see fig. 13), but instead improved K for colours $F(1, 10) = 7.17$, $p = .023$; $\eta_p^2 = .269$ (see fig. 14) while having no influence on C $F(1, 10) = 0.263$, $p = .619$; $\eta_p^2 = .004$ and t_0 $F(1, 10) = 0.052$, $p = .825$; $\eta_p^2 = .001$ (see fig. 15 – 16).

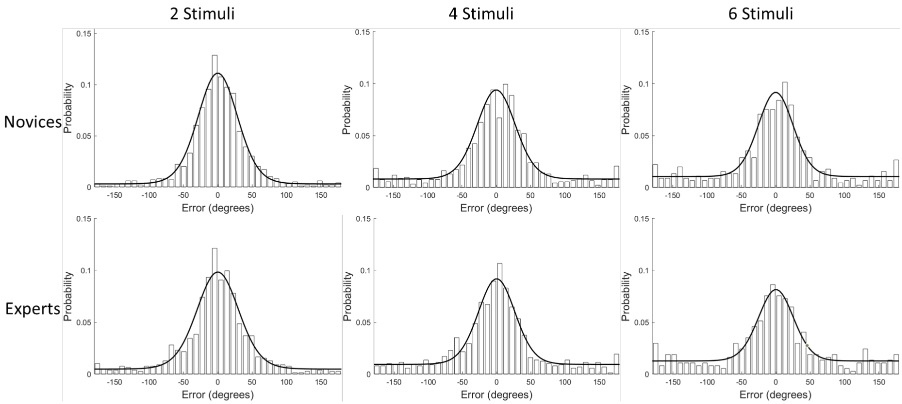

Fig. 13 The aggregate MemToolbox models from paper 3 across all participants for each set size (two, four, and six) across both groups (Novices and Experts).

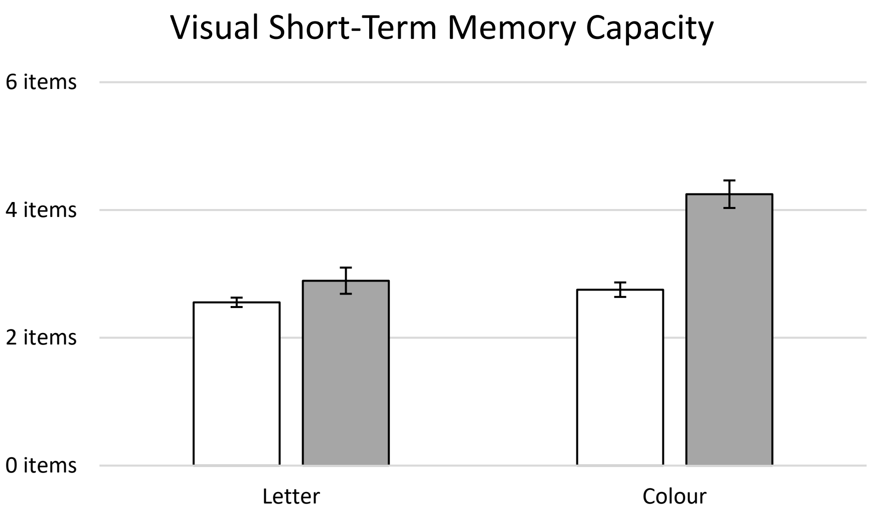

Fig. 14 The visual short-term memory capacity for both groups (Novices in white and Experts in black) and conditions (Letter and Colour). Error bars denote 95% confidence interval.

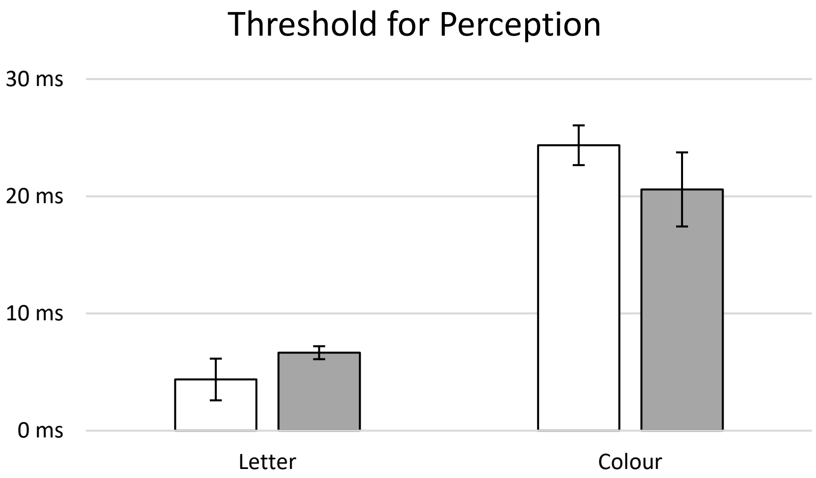

Fig. 15 The threshold for perception for both groups (Novices in white and Experts in black) and conditions (Letter and Colour). Error bars denote 95% confidence interval.

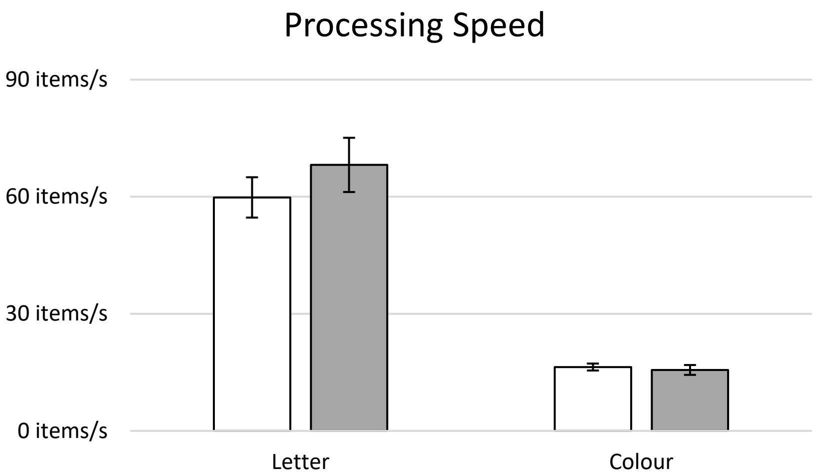

Fig. 16 The processing speed for both groups (Novices in white and Experts in black) and conditions (Letter and Colours). Error bars denote 95% confidence interval.

The lack of any expertise-related *C* differences suggests at least some level of colour expertise for the novices, as expertise has an influence on low expertise items (Dall et al., 2021; Pelli Burns, Farell, & Moore-Page, 2006; Xie & Zhang, 2017), while having no effect on high expertise items (Eng, Chen, & Jiang, 2005; Pelli, et al., 2006). The large-stimuli specific *K* effect is in line with previous studies that found an expertise related *K* improvement (e.g. Dall et al., 2021; Jackson & Raymond, 2005; Sun et al., 2011; Sørensen & Kyllingsbæk, 2012; Xie & Zhang, 2017; 2018). This difference might reflect differences in perceptual strategies due to differences in mental categories (Brogaard & Sørensen, *in press*), such as investigations that report improved discriminability of shades of blue for Greek individuals, who have more words for the colour (e.g. Athanasopoulos, 2009; Thierry et al., 2009). Our *K* improvements in combination with the apparent lack of changes in precision contradict previous studies conducted using face stimuli (Lorenc et al., 2014; Scolari et al., 2008). It is, however, similar to a pattern described by Ye and colleagues (2019), who found that high memory capacity increased the number of remembered items when they were shown at 50ms/item, but increased the precision when shown at 100ms/item. This suggests that expertise improves memory capacity when there is limited time to encode, but increases the precision of stored items when participants have time to encode it all.

CHAPTER 4. DISCUSSION

The aim of this thesis is to investigate the influence of prior information on VSTM. With that aim in mind, three investigations were conducted with varying levels of stimuli specific expertise in the participants. On the basis of our empirical data, the thesis has extended the expertise effects observed by several authors (e.g. Curby et al., 2009; Jackson & Raymond, 2005; Sun et al., 2011; Sørensen & Kyllingsbæk, 2012; Xie & Zhang, 2017b) in several change detection or resolution based paradigms, to whole report using discreet and continuous stimuli that participants might not know by name. All papers (paper 1 – 3) used whole report paradigms where the participants chose what stimuli was shown from a range of options, and they all demonstrate similar findings. This is important, as Donkin and colleagues (2016) suggest that different designs might show different results. The empirical data were analysed in a similar vein with TVA, which allowed the isolation of specific cognitive parameters, with a specific focus on K, C, and t_0. All three papers (paper 1 – 3) either showed or indicated an expertise related K improvement with no significant difference in t_0. Paper 1 specifically showed how different levels of Japanese expertise improved the number of Japanese hiragana participants were able to remember, while failing to show any significant difference in the number of remembered Latin letters or simple picture drawings. More specifically, participants lacking any Japanese expertise were able to remember fewer hiragana than those participants who were studying Japanese at university and native speakers, whereas native speakers could remember more than participants were studying Japanese at university. This increase seems to come early in training as participants who had studied Japanese showed significantly higher K for hiragana.

However, several authors have had great difficulty in showing K improvements as a result of training (Chen, Eng, & Jiang 2006; Olson & Jiang, 2004), though these participants received comparatively little training compared to the participants in Dall and colleagues (2016). In paper 2, much the same as in paper 1, we found that there seems to be an influence of expertise within languages. We did this by examining how the frequency of use of a Chinese character and the stroke count of these characters influenced VSTM in native Chinese speakers. We found that participants were able to remember more commonly used characters compared to uncommon ones. What initially seems striking is that there was no influence of stroke count, meaning that uncommon characters with a low stroke count (e.g. 匕) were more difficult to remember than a common one with a high stroke count (e.g. 辰) (Dall et al., 2021). This is especially interesting in the light of previous findings that stroke count had an influence for participants who did not speak Chinese (e.g. Sun, et al., 2011). However, Dall and colleagues (2021) suggested the lack of influence of stroke count might be caused by Chinese university students having some degree of expertise even for fairly uncommon Chinese characters.

It is worth noting that increased *K* may not be reflective of an actual increase in VSTM, but rather an increase in the mental resolution of the stored stimuli (Scolari, Vogel, & Awh, 2008). As a result, the final paper combined TVA with the MemToolbox (Suchow, Brady, Fougnie, & Alvarez, 2013) in order to study the influence of stimulus specific expertise in relation to continuous categories. We did this by showing participants either Latin letters or coloured circles, with the letters serving as a control condition. Participants as colour experts were students at the School of Visual Arts from the Royal Danish Academy of Fine Arts, and the colour novices were psychology students with no formal training in painting, graphical design or other visual arts. Just as in the previous two studies, we found a significant expertise related *K* improvement for experts compared to the novices for colours, with no significant difference in *C* and t_0. The lack of expertise related *C* improvement could be explained by a combination of two things: the existing level of colour expertise in the novices and the lack of important features in the coloured circles. There was no significant group difference in the precision of reported colour stimuli. This *K* improvement with no precision difference seems to run counter to several studies showing improved precision with no *K* improvements (Scolari et al., 2008; Lorence et al., 2014). It should be noted, however, that expertise was made by inverting pictures in both Scolari and colleagues (2008) and Lorence and colleagues (2014). A study by Zhou, Mondloch, and Emrich (2018) used the other-race effect in a circular face space (see fig. 17) to model expertise and found no significant expertise related precision differences, while finding changes in the memory capacity.

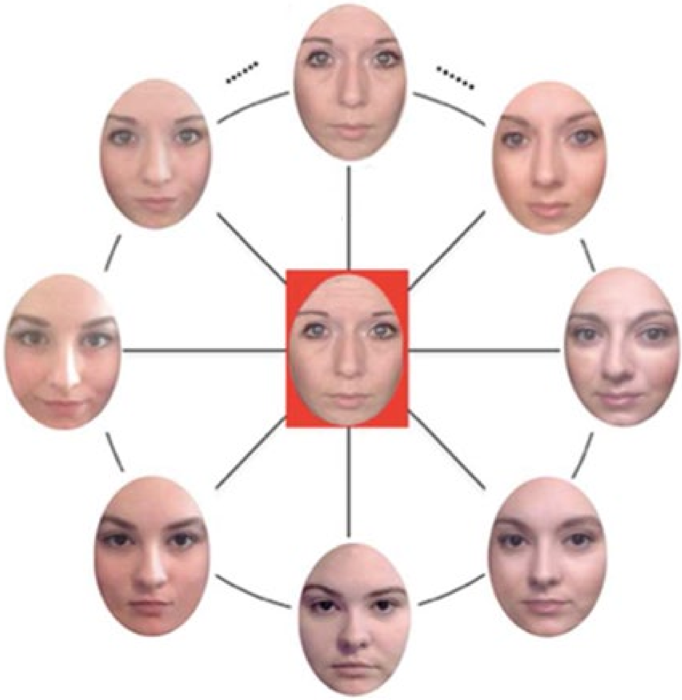

Fig. 17. Continuous face space stimuli used in Zhou and colleagues (2018).

Oberauer and Eichenberger (2013) discussed different types of bottlenecks in relation to VSTM, and concluded that there were three types in optimal viewing conditions: number of objects, number of relevant features per object, and precision of memory for each feature. Their arguments in favour of the influence of the number of features initially seem to contradict the lack of stroke count influence found in Dall and colleagues (2021) (paper 2). However, it is important to note the difference in the type of stimuli used in the two studies and the use of features. Oberauer and Eichenberger (2013) defined features by the number of different types of feature that need to be held in VSTM (i.e. the colour, size, and gradient of squares) while Dall and colleagues (2021) defined features as how many features the single object was composed of (i.e. the number of stroke count in Chinese characters). This difference might explain why Dall and colleagues (2021) found no influence of stroke count as participants might have been able to encode the different stimuli as single feature stimulus based on the smallest discriminable feature (Swan et al., 2016). Paper 3, furthermore, would suggest that it is not possible to explain the VSTM bottleneck solely using the three different types of bottlenecks proposed by Oberauer and Eichenberger (2013) as it demonstrated that K in the VSTM bottleneck can be increased independent of the number of objects, relevant features per object and the precision of each feature.

Alternatively, Paper 3 may indicate that expertise is more related to processing and maintaining the smallest discriminable parts of the target stimuli (Swan, Collins, & Wyble, 2016). It is however important to note the difference between maintaining and processing, as the smallest discriminable difference for maintaining two items might not be a visual component. An example could be the Chinese character for hand (手) and head (首) which have the same pronunciation (Shou). This thesis focusses on visual processing rather than maintenance. Processing by locating the smallest discriminable part can be likened to a digital signal consisting of a series of bits of information (Shannon, 1948), with each bit having one neural activation (Hubel & Wiesel, 1959). Bundesen and colleagues (2011) proposed that visual attention proceeds through a recurring feedback loop between PFC and a number of mental templates. Sørensen and Kyllingsbæk (2012) suggest that expertise strengthens the feedback loops to these mental templates. This compression of visual objects into mental categories explains why deviation from templates, e.g. 'L' from '*L*', would reduce the efficiency of processing (Price, McElroy and Martin, 2016).

The question remains of how these templates are created and of their limitations. What is the difference between discrete and continuous categories? Are continuous categories 'just' several discrete categories, as suggested by Brody and colleagues (2003), or are they different? When does a distorted discrete stimuli change from being in one category to another, and is this change based on the strength of the original categories? This merits a tougher examination of the bottlenecks of η in future research.

CHAPTER 5. CONCLUSION

Memory and attention research have come a long way since the early days of the cognitive revolution. However, as methodology has improved, we have found an increasing number of factors and bottlenecks that constrain VSTM. Prior expertise is shown to be highly involved in memory, to the point where the visual complexity of an item might be negated by stimuli specific expertise (e.g. Dall et al., 2021; Sun et al., 2011). The question still remains as to how this expertise develops, and at what point visual complexity disappears.

CHAPTER 6. REFERENCE LIST

1. Adam, K. C., Vogel, E. K., & Awh, E. (2017). Clear evidence for item limits in visual working memory. *Cognitive Psychology*, *97*, 79-97. https://doi.org/10.1016/j.cogpsych.2017.07.001
2. Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological science*, *15*(2), 106-111. https://doi.org/10.1111/j.0963-7214.2004.01502006.x
3. Atick, J. J., & Redlich, A. N. (1990). Towards a theory of early visual processing. *Neural computation*, *2*(3), 308-320. https://doi.org/10.1162/neco.1990.2.3.308
4. Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In *Psychology of learning and motivation* (Vol. 2, pp. 89-195). Academic Press. https://doi.org/10.1016/S0079-7421(08)60422-3
5. Averbach, E., & Coriell, A. S. (1961). Short-term memory in vision. *The Bell System Technical Journal*, *40*(1), 309-328. https://doi.org/10.1002/j.1538-7305.1961.tb03987.x
6. Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47-89). Academic press. https://doi.org/10.1016/S0079-7421(08)60452-1

7. Baddeley, A., Lewis, V., & Vallar, G. (1984). Exploring the articulatory loop. *The Quarterly Journal of Experimental Psychology Section A*, *36*(2), 233-252. https://doi.org/10.1080/14640748408402157

8. Baker, H. D. (1963). Initial stages of dark and light adaptation. *JOSA*, *53*(1), 98-103. https://doi.org/10.1364/JOSA.53.000098

9. Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, *321*(5890), 851-854. https://doi.org/10.1126/science.1158023

10. Bays, P. M., Catalao, R. F., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of vision*, *9*(10), 7-7. https://doi.org/10.1167/9.10.7

11. Baxt, N. (1871). Ueber die Zeit, welche nöthig ist, damit ein Gesichtseindruck zum Bewusstsein kommt und über die Grösse (Extension) der bewussten Wahrnehmung bei einem Gesichtseindrucke von gegebener Dauer. *Pflügers Archiv European Journal of Physiology*, *4*(1), 325-336. https://doi.org/10.1007/BF01612494

12. Baxt, N. (1982). On the time necessary for a visual impression to come into consciousness. *Psychological Research*, *44*(1), 1-12. https://doi.org/10.1007/BF00308551

13. Blalock, L. D. (2013). Mask similarity impacts short-term consolidation in visual working memory. *Psychonomic bulletin & review*, *20*(6), 1290-1295. https://doi.org/10.3758/s13423-013-0461-9

14. Box, G. E. (1976). Science and statistics. *Journal of the American Statistical Association*, *71*(356), 791-799. https://doi.org/10.1080/01621459.1976.10480949

15. Box, G. E. (1979). Robustness in the strategy of scientific model building. In *Robustness in statistics* (pp. 201-236). Academic Press. https://doi.org/10.1016/B978-0-12-438150-6.50018-2

16. Boynton, R. M., & Kandel, G. (1957). On responses in the human visual system as a function of adaptation level. *JOSA*, *47*(4), 275-286. https://doi.org/10.1364/JOSA.47.000275

17. Breitmeyer, B. G., & Ganz, L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. *Psychological review*, *83*(1), 1-36. https://doi.org/10.1037/0033-295X.83.1.1

18. Broadbent, D. E. (1958). *Perception and communication*. New York: Oxford University Press. https://doi.org/10.1016/c2013-0-08164-9

19. Brody, C. D., Romo, R., & Kepecs, A. (2003). Basic mechanisms for graded persistent activity: discrete attractors, continuous attractors, and dynamic representations. *Current Opinion in Neurobiology*, *13*(2), 204-211. https://doi.org/10.1016/S0959-4388(03)00050-3

20. Bundesen, C. (1987). Visual attention: Race models for selection from multielement displays. *Psychological Research*, *49*(2-3), 113-121. https://doi.org/10.1007/BF00308676

21. Bundesen, C. (1990). A theory of visual attention. *Psychological review*, *97*(4), 523-547. https://doi.org/10.1037/0033-295X.97.4.523
22. Bundesen, C. & Habekost, T. (2008). *Principles of Visual Attention: Linking Mind and Brain*. Oxford University Press Oxford. https://doi.org/10.1093/acprof:oso/9780198570707.001.0001
23. Bundesen, C., Vangkilde, S., & Habekost, T. (2015). Components of visual bias: a multiplicative hypothesis. *Annals of the New York Academy of Sciences*, *1339*(1), 116-124. https://doi.org/10.1111/nyas.12665
24. Buttle, H., & Raymond, J. E. (2003). High familiarity enhances visual change detection for face stimuli. *Perception & psychophysics*, *65*(8), 1296-1306. https://doi.org/10.3758/BF03194853
25. Caspersen, I. D., Petersen, A., Vangkilde, S., Plessen, K. J., & Habekost, T. (2017). Perceptual and response-dependent profiles of attention in children with ADHD. *Neuropsychology*, *31*(4), 349-360. http://dx.doi.org/10.1037/neu0000375
26. Cattell, J. M. (1885). Über die Zeit der Erkennung und Benennung von Schriftzeichen, Bildern und Farben. *Philosophische Studien*, *2*, 635-650. https://doi.org/10.2307/1410904
27. Cattell, J. M. (1890). Mental tests and measurements. *Mind*, *15*(59), 373-380.
28. Chadwick, M. J., Anjum, R. S., Kumaran, D., Schacter, D. L., Spiers, H. J., & Hassabis, D. (2016). Semantic representations in the temporal pole predict

false memories. *Proceedings of the National Academy of Sciences*, *113*(36), 10180-10185. https://doi.org/10.1073/pnas.1610686113

29. Christensen, J. H., Markussen, B., Bundesen, C., & Kyllingsbæk, S. (2018). A physiologically based nonhomogeneous Poisson counter model of visual identification. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(9), 1383-1398. https://doi.org/10.1037/xhp0000539
30. Coltheart, M., & Arthur, B. (1972). Evidence for an integration theory of visual masking. *Quarterly Journal of Experimental Psychology*, *24*(3), 262-269. https://doi.org/10.1080/14640747208400279
31. Conway, A. R., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, *30*(2), 163-183. https://doi.org/10.1016/S0160-2896(01)00096-4
32. Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic bulletin & review*, *12*(5), 769-786. https://doi.org/10.3758/BF03196772
33. Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and brain sciences*, *24*(1), 87-114. https://doi.org/10.1017/S0140525X01003922
34. Cowan, N. (2015). George Miller's magical number of immediate memory in retrospect: Observations on the faltering progression of

science. *Psychological review*, *122*(3), 536-541. http://dx.doi.org/10.1037/a0039035

35. Crawford, B. H. (1947). Visual adaptation in relation to brief conditioning stimuli. *Proceedings of the Royal Society of London. Series B-Biological Sciences*, *134*(875), 283-302. https://doi.org/10.1098/rspb.1947.0015
36. Crouzet, S. M., Overgaard, M., & Busch, N. A. (2014). The fastest saccadic responses escape visual masking. *PloS one*, *9*(2), e87418, 1-6. https://doi.org/10.1371/journal.pone.0087418
37. Crystal, D. *The Cambridge encyclopedia of language* (Vol. 1988). Cambridge: Cambridge University Press (1987).
38. Curby, K. M., & Gauthier, I. (2007). A visual short-term memory advantage for faces. *Psychonomic bulletin & review*, *14*(4), 620-628. https://doi.org/10.3758/BF03196811
39. Curby, K. M., Glazek, K., & Gauthier, I. (2009). A visual short-term memory advantage for objects of expertise. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), 94-107. https://doi.org/10.1037/0096-1523.35.1.94
40. Da, J. A corpus-based study of character and bigram frequencies in Chinese e-texts and its implications for Chinese language instruction. In *Proceedings of the fourth international conference on new technologies in teaching and learning Chinese* (501-511). Beijing: Tsinghua University Press (2004).
41. Dall, J. O., Wang, Y. M., Cai, X. L., Chan, R. C., & Sørensen, T. A. (2021). Visual short-term memory and attention: An investigation of

familiarity and stroke count in Chinese characters. *Journal of Experimental Psychology. Learning, Memory, and Cognition*. *47*(2), 282–294. https://doi.org/10.1037/xlm0000950

42. Dall, J. O., Watanabe, K., & Sørensen, T. A. (2016). Category specific knowledge modulate capacity limitations of visual short-term memory. In *8th International Conference on Knowledge and Smart Technology (KST)*, (275-280). IEEE. http://dx.doi.org/10.1109/KST.2016.7440508
43. Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior*, *19*(4), 450-466. https://doi.org/10.1016/S0022-5371(80)90312-6
44. Das, S., & Cakmak, U. M. (2018). *Hands-On Automated Machine Learning: A beginner's guide to building automated machine learning systems using AutoML and Python*. Packt Publishing Ltd.
45. Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: the psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General*, *129*(4), 481-507. https://doi.org/10.1037/0096-3445.129.4.481
46. Donders, F. C. (1969). On the speed of mental processes. *Acta psychologica*, *30*, 412-431. https://doi.org/10.1016/0001-6918(69)90065-1
47. Donkin, C., Kary, A., Tahir, F., & Taylor, R. (2016). Resources masquerading as slots: Flexible allocation of visual working memory. *Cognitive Psychology*, *85*, 30-42. https://doi.org/10.1016/j.cogpsych.2016.01.002

48. Donkin, C., Tran, S. C., & Nosofsky, R. (2014). Landscaping analyses of the ROC predictions of discrete-slots and signal-detection models of visual working memory. *Attention, Perception, & Psychophysics*, *76*(7), 2103-2116. https://doi.org/10.3758/s13414-013-0561-7

49. Duncan, J., Bundesen, C., Olson, A., Humphreys, G., Chavda, S., & Shibuya, H. (1999). Systematic analysis of deficits in visual attention. *Journal of Experimental Psychology: General*, *128*(4), 450-478. https://doi.org/10.1037/0096-3445.128.4.450

50. Dyrholm, M., Kyllingsbæk, S., Espeseth, T., & Bundesen, C. (2011). Generalizing parametric models by introducing trial-by-trial parameter variability: The case of TVA. *Journal of Mathematical Psychology*, *55*(6), 416-429. https://doi.org/10.1016/j.jmp.2011.08.005

51. Ericsson, K. A., & Chase, W. G. (1982). Exceptional memory: Extraordinary feats of memory can be matched or surpassed by people with average memories that have been improved by training. *American Scientist*, *70*(6), 607-615. Retrieved from http://www.jstor.org/stable/27851732

52. Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological review*, *102*(2), 211-245. https://doi.org/10.1037/0033-295X.102.2.211

53. Everitt, B. S., Palmer, C. R. (eds) (2005). *Encyclopaedic companion to medical statistics*, London: Arnold.

54. Finke, K., Dodds, C. M., Bublak, P., Regenthal, R., Baumann, F., Manly, T., & Müller, U. (2010). Effects of modafinil and methylphenidate on visual

attention capacity: a TVA-based study. *Psychopharmacology*, *210*(3), 317-329. https://doi.org/10.1007/s00213-010-1823-x

55. Frick, R. W. (1988). Issues of representation and limited capacity in the visuospatial sketchpad. *British Journal of Psychology*, *79*(3), 289-308. https://doi.org/10.1111/j.2044-8295.1988.tb02289.x
56. Gao, Z., Yin, J., Xu, H., Shui, R., & Shen, M. (2011). Tracking object number or information load in visual working memory: Revisiting the cognitive implication of contralateral delay activity. *Biological Psychology*, *87*(2), 296-302. https://doi.org/10.1016/j.biopsycho.2011.03.013
57. Gauthier, I., Behrmann, M., & Tarr, M. J. (2004). Are Greebles like faces? Using the neuropsychological exception to test the rule. *Neuropsychologia*, *42*(14), 1961-1970. https://doi.org/10.1016/j.neuropsychologia.2004.04.025
58. Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature neuroscience*, *3*(2), 191-197. https://doi.org/10.1038/72140
59. Gauthier, I., & Tarr, M. J. (1997). Becoming a “Greeble” expert: Exploring mechanisms for face recognition. *Vision research*, *37*(12), 1673-1682. https://doi.org/10.1016/S0042-6989(96)00286-6
60. Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform 'face area' increases with expertise

in recognizing novel objects. *Nature neuroscience*, *2*(6), 568-573. https://doi.org/10.1038/9224

61. Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature neuroscience*, *7*(5), 555-562. https://doi.org/10.1038/nn1224
62. Gobet, F., & Simon, H. A. (1996). Templates in chess memory: A mechanism for recalling several boards. *Cognitive psychology*, *31*(1), 1-40. https://doi.org/10.1006/cogp.1996.0011
63. Grelotti, D. J., Klin, A. J., Gauthier, I., Skudlarski, P., Cohen, D. J., Gore, J. C., Volkmar, F. R., & Schultz, R. T. (2005). fMRI activation of the fusiform gyrus and amygdala to cartoon characters but not to faces in a boy with autism. *Neuropsychologia*, *43*(3), 373-385. https://doi.org/10.1016/j.neuropsychologia.2004.06.015
64. Gross, C. G., Rocha-Miranda, C. D., & Bender, D. B. (1972). Visual properties of neurons in inferotemporal cortex of the macaque. *Journal of neurophysiology*, *35*(1), 96-111. https://doi.org/10.1152/jn.1972.35.1.96
65. Gögler, N., Willacker, L., Funk, J., Strube, W., Langgartner, S., Napiórkowski, N., Hasan, A., & Finke, K. (2016). Single-session transcranial direct current stimulation induces enduring enhancement of visual processing speed in patients with major depression. *European Archives of Psychiatry and Clinical Neuroscience*, 1-16. https://doi.org/10.1007/s00406-016-0761-y

66. Habekost, T. (2015). Clinical TVA-based studies: a general review. *Frontiers in Psychology*, *6*, 290, 1-18. https://doi.org/10.3389/fpsyg.2015.00290

67. Habekost, T., & Bundesen, C. (2003). Patient assessment based on a theory of visual attention (TVA): Subtle deficits after a right frontal-subcortical lesion. *Neuropsychologia*, *41*(9), 1171-1188. https://doi.org/10.1016/S0028-3932(03)00018-6

68. Hebb, D. O. (1949). *The organization of behaviour; a neuropsychological theory*. Oxford, England: Wiley.

69. Hilla, Y., von Mankowski, J., Föcker, J., & Sauseng, P. (2020). Faster Visual Information Processing in Video Gamers Is Associated With EEG Alpha Amplitude Modulation. *Frontiers in Psychology*, *11*, 3333, 1-16. https://doi.org/10.3389/fpsyg.2020.599788

70. Hintzman, D. L. (1965). Classification and aural coding in short-term memory. *Psychonomic Science*, *3*(1-12), 161-162. https://doi.org/10.3758/BF03343073

71. Hirose, N., & Osaka, N. (2009). Object substitution masking induced by illusory masks: Evidence for higher object-level locus of interference. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(4), 931-938. https://doi.org/10.1037/a0012734

72. Huang, L. (2010). Visual working memory is better characterized as a distributed resource rather than discrete slots. *Journal of Vision*, *10*(14), 8-8. https://doi.org/10.1167/10.14.8

73. Hubel, D. H. (1959). Single unit activity in striate cortex of unrestrained cats. *The Journal of physiology*, *147*(2), 226-238. https://doi.org/10.1113/jphysiol.1959.sp006238
74. Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *The Journal of physiology*, *148*(3), 574-591. https://doi.org/10.1113/jphysiol.1959.sp006308
75. Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology*, *160*(1), 106-154. https://doi.org/10.1113/jphysiol.1962.sp006837
76. Hubel, D. H., & Wiesel, T. N. (1963). Shape and arrangement of columns in cat's striate cortex. *The Journal of physiology*, *165*(3), 559-568. https://doi.org/10.1113/jphysiol.1963.sp007079
77. Hubel, D. H., & Wiesel, T. N. (1965). Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *Journal of neurophysiology*, *28*(2), 229-289. https://doi.org/10.1152/jn.1965.28.2.229
78. Hulme, C., Maughan, S., & Brown, G. D. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of memory and language*, *30*(6), 685-701. https://doi.org/10.1016/0749-596X(91)90032-F
79. Hulme, C., Roodenrys, S., Brown, G., & Mercer, R. (1995). The role of long-term memory mechanisms in memory span. *British Journal of*

Psychology, *86*(4), 527-536. https://doi.org/10.1111/j.2044-8295.1995.tb02570.x

80. Igarashi, Y. (2007). *The changing role of katakana in the Japanese writing system* (Doctoral dissertation).
81. Jackson, M. C., Linden, D. E., Roberts, M. V., Kriegeskorte, N., & Haenschel, C. (2015). Similarity, not complexity, determines visual working memory performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(6), 1884-1892. https://doi.org/10.1037/xlm0000125
82. Jackson, M. C., & Raymond, J. E. (2005). Visual Working Memory: capacity is dependent on perceived, not physical, stimulus complexity. *Journal of Vision*, **5**(8)**,** 621-621. https://doi.org/10.1167/5.8.621
83. Jacobs, J. (1887). Experiments on "Prehension". *Mind*, *12*(45), 75-79. Retrieved from http://www.jstor.org/stable/2246990
84. James, W. (1890). *The principles of psychology, Vol. 1*. Henry Holt and Co. https://doi.org/10.1037/10538-000
85. James, T. W., & James, K. H. (2013). Expert individuation of objects increases activation in the fusiform face area of children. *NeuroImage*, *67*, 182-192. https://doi.org/10.1016/j.neuroimage.2012.11.007
86. Kahneman, D. (1968). Method, findings, and theory in studies of visual masking. *Psychological Bulletin*, *70*(6p1), 404-425. https://doi.org/10.1037/h0026731

87. Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Englewood Cliffs, NJ: Prentice-Hall.

88. Kane, M. J., Conway, A. R., Miura, T. K., & Colflesh, G. J. (2007). Working memory, attention control, and the N-back task: a question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 615-622. https://doi.org/10.1037/0278-7393.33.3.615

89. Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of neuroscience*, *17*(11), 4302-4311. https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997

90. Kinsbourne, M., & Warrington, E. K. (1962). Further studies on the masking of brief visual stimuli by a random pattern. *Quarterly Journal of Experimental Psychology*, *14*(4), 235-245. https://doi.org/10.1080/17470216208416541

91. Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain research reviews*, *29*(2-3), 169-195. https://doi.org/10.1016/S0165-0173(98)00056-3

92. Klimesch, W., Doppelmayr, M., Schimke, H., & Pachinger, T. (1996). Alpha frequency, reaction time, and the speed of processing information. *Journal of clinical neurophysiology*, *13*(6), 511-518. https://doi.org/10.1097/00004691-199611000-00006

93. Klimesch, W., Schack, B., Schabus, M., Doppelmayr, M., Gruber, W., & Sauseng, P. (2004). Phase-locked alpha and theta oscillations generate the

P1–N1 complex and are related to memory performance. *Cognitive Brain Research*, *19*(3), 302-316. https://doi.org/10.1016/j.cogbrainres.2003.11.016

94. Kraft, A., Dyrholm, M., Bundesen, C., Kyllingsbæk, S., Kathmann, N., & Brandt, S. A. (2013). Visual attention capacity parameters covary with hemifield alignment. *Neuropsychologia*, *51*(5), 876-885. https://doi.org/10.1016/j.neuropsychologia.2013.01.021
95. Krüger, A., Tünnermann, J., & Scharlau, I. (2017). Measuring and modeling salience with the theory of visual attention. *Attention, Perception, & Psychophysics*, *79*(6), 1593-1614. https://doi.org/10.3758/s13414-017-1325-6
96. Kyllingsbæk, S. (2006). Modeling visual attention. *Behavior Research Methods*, *38*(1), 123-133. https://doi.org/10.3758/BF03192757
97. Kyllingsbæk, S., Markussen, B., & Bundesen, C. (2012). Testing a Poisson counter model for visual identification of briefly presented, mutually confusable single stimuli in pure accuracy tasks. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(3), 628-642. https://doi.org/10.1037/a0024751
98. Li, A. Y., Liang, J. C., Lee, A. C. H., & Barense, M. D. (2020). The validated circular shape space: Quantifying the visual similarity of shape. *Journal of Experimental Psychology: General, 149*(5), 949–966. https://doi.org/10.1037/xge0000693

99. Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279-281. https://doi.org/10.1038/36846

100. Marr, D. (1976). Early processing of visual information. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *275*(942), 483-519. https://doi.org/10.1098/rstb.1976.0090

101. McElree, B. (2001). Working memory and focal attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(3), 817-835. https://doi.org/10.1037/0278-7393.27.3.817

102. McGugin, R. W., Newton, A. T., Gore, J. C., & Gauthier, I. (2014). Robust expertise effects in right FFA. *Neuropsychologia*, *63*, 135-144. https://doi.org/10.1016/j.neuropsychologia.2014.08.029

103. Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, *63*(2), 81-97. https://doi.org/10.1037/h0043158

104. Miller, G. A. (2003). The cognitive revolution: a historical perspective. *Trends in cognitive sciences*, *7*(3), 141-144. https://doi.org/10.1016/S1364-6613(03)00029-9

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

106. Oberauer, K., & Eichenberger, S. (2013). Visual working memory declines when more features must be remembered for each object. *Memory & cognition*, *41*(8), 1212-1227. https://doi.org/10.3758/s13421-013-0333-6

107. Ojemann, J. G., Ojemann, G. A., & Lettich, E. (1992). Neuronal activity related to faces and matching in human right nondominant temporal cortex. *Brain*, *115*(1), 1-13. https://doi.org/10.1093/brain/115.1.1

108. Olson, I. R., & Jiang, Y. (2004). Visual short-term memory is not improved by training. *Memory & cognition*, *32*(8), 1326-1332. https://doi.org/10.3758/BF03206323

109. Olsson, H., & Poom, L. (2005). Visual memory needs categories. *Proceedings of the National Academy of Sciences*, *102*(24), 8776-8780. https://doi.org/10.1073/pnas.0500810102

110. Pashler, H. (1988). Familiarity and visual change detection. *Perception & psychophysics*, *44*(4), 369-378. https://doi.org/10.3758/BF03210419

111. Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision research*, *46*(28), 4646-4674. https://doi.org/10.1016/j.visres.2006.04.023

112. Petersen, A., & Andersen, T. S. (2012). The effect of exposure duration on visual character identification in single, whole, and partial report. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(2), 498-514. https://doi.org/10.1037/a0026728

113. Peterson, L. R. (1969). Concurrent verbal activity. *Psychological Review*, *76*(4), 376-386. https://doi.org/10.1037/h0027443

114. Phillips, D. L. (1962). A technique for the numerical solution of certain integral equations of the first kind. *Journal of the ACM (JACM)*, *9*(1), 84-97. https://doi.org/10.1145/321105.321114

115. Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics*, *16*(2), 283-290. https://doi.org/10.3758/BF03203943

116. Piéron, H. (1925). I. Recherches expérimentales sur la marge de variation du temps de latence de la sensation lumineuse (par une méthode de masquage). *L'année psychologique*, *26*(1), 1-30.

117. Price, J., McElroy, K., & Martin, N. J. (2016). The role of font size and font style in younger and older adults' predicted and actual recall performance. *Aging, Neuropsychology, and Cognition*, *23*(3), 366-388. https://doi.org/10.1080/13825585.2015.1102194

118. Purves, D., Augustine, G. J., Fitzpatrick, D., Hall, W. C., LaMantia A., Mooney, R. D., Platt, M. L., & White, L. E., (2018). *Neuroscience* (6th ed.). Sunderland, Massachusetts: Oxford University Press.

119. Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia--a window into perception, thought and language. *Journal of consciousness studies*, *8*(12), 3-34. https://doi.org/10.1111/1468-0068.00363

120. Ramsøy, T. Z., & Overgaard, M. (2004). Introspection and subliminal perception. *Phenomenology and the cognitive sciences*, *3*(1), 1-23. https://doi.org/10.1023/B:PHEN.0000041900.30172.e8

121. Richardson, J. T. E., & Baddeley, A. D. (1975). The effect of articulatory suppression in free recall. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 623-629. https://doi.org/10.1016/S0022-5371(75)80049-1

122. Rossion, B., Kung, C. C., & Tarr, M. J. (2004). Visual expertise with nonface objects leads to competition with the early perceptual processing of faces in the human occipitotemporal cortex. *Proceedings of the National Academy of Sciences*, *101*(40), 14521-14526. https://doi.org/10.1073/pnas.0405613101

123. Santosa, F., & Symes, W. W. (1986). Linear inversion of band-limited reflection seismograms. *SIAM Journal on Scientific and Statistical Computing*, *7*(4), 1307-1330. https://doi.org/10.1137/0907087

124. Schubert, T., Finke, K., Redel, P., Kluckow, S., Müller, H., & Strobach, T. (2015). Video game experience and its influence on visual attention parameters: an investigation using the framework of the Theory of Visual Attention (TVA). *Acta psychologica*, *157*, 200-214. https://doi.org/10.1016/j.actpsy.2015.03.005

125. Scolari, M., Vogel, E. K., & Awh, E. (2008). Perceptual expertise enhances the resolution but not the number of representations in working memory. *Psychonomic bulletin & review*, *15*(1), 215-222. https://doi.org/10.3758/PBR.15.1.215

126. Shannon, C. E. (1948). A mathematical theory of communication. *Bell system technical journal*, *27*(3), 379-423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

127. Shibuya, H., & Bundesen, C. (1988). Visual selection from multielement displays: measuring and modeling effects of exposure duration. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(4), 591-600. https://doi.org/10.1037/0096-1523.14.4.591

128. Sligte, I. G., Scholte, H. S., & Lamme, V. A. (2008). Are there multiple visual short-term memory stores?. *PLOS one*, *3*(2), e1699, 1-9. https://doi.org/10.1371/journal.pone.0001699

129. Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of experimental psychology: Human learning and memory*, *6*(2), 174-215. https://doi.org/10.1037/0278-7393.6.2.174

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

131. Sperling, G. (1960b). Negative afterimage without prior positive image. *Science*, *131*(3413), 1613-1614. https://doi.org/10.1126/science.131.3413.1613

132. Sperling, G. (1963). A model for visual memory tasks. *Human factors*, *5*(1), 19-31. https://doi.org/10.1177/001872086300500103

133. Sperling, G. (1964). What visual masking can tell us about temporal factors in perception. In *Proceedings of the seventeenth International Congress of Psychology* (pp. 199-200). Amsterdam: North-Holland.

134. Sperling, G. (1967). Successive approximations to a model for short term memory. *Acta psychologica*, *27*, 285-292. https://doi.org/10.1016/0001-6918(67)90070-4

135. Stigler, R. (1910). Chronophotische Studien über den Umgebungskontrast. *Pflügers Archiv European Journal of Physiology*, *134*(6), 365-435. https://doi.org/10.1007/BF01680352

136. Suchow, J. W., Brady, T. F., Fougnie, D., & Alvarez, G. A. (2013). Modeling visual working memory with the MemToolbox. *Journal of Vision*, *13*(10), 1-8. https://doi.org/10.1167/13.10.9

137.

138. Sun, H., Zimmer, H. D., & Fu, X. (2011). The influence of expertise and of physical complexity on visual short-term memory consolidation. *The Quarterly Journal of Experimental Psychology*, *64*(4), 707-729. https://doi.org/10.1080/17470218.2010.511238

139. Swan, G., Collins, J., & Wyble, B. (2016). Memory for a single object has differently variable precisions for relevant and irrelevant features. *Journal of vision*, *16*(3), 1-12. https://doi.org/10.1167/16.3.32

140. Sørensen, T. A., & Kyllingsbæk, S. (2009). The nature of the capacity limitations in visual short-term memory. *Journal of Vision*, *9*(8), 589-589. https://doi.org/10.1167/9.8.589

141. Sørensen, T. A., & Kyllingsbæk, S. (2012). Short-term storage capacity for visual objects depends on expertise. *Acta psychologica*, *140*(2), 158-163. https://doi.org/10.1016/j.actpsy.2012.04.004

142. Sørensen, T. A., Vangkilde, S., & Bundesen, C. (2015). Components of attention modulated by temporal expectation. *Journal of experimental psychology: learning, memory, and cognition*, *41*(1), 178-192. https://doi.org/10.1037/a0037268

143. Taft, M., & Zhu, X. (1997). Submorphemic processing in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(3), 761–775. https://doi.org/10.1037/0278-7393.23.3.761

144. Tarr, M. J., & Gauthier, I. (2000). FFA: a flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nature neuroscience*, *3*(8), 764-769. https://doi.org/10.1038/77666

145. Theeuwes, J. (1991). Cross-dimensional perceptual selectivity. *Perception & psychophysics*, *50*(2), 184-193. https://doi.org/10.3758/BF03212219

146. Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267-288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

147. Tikhonov, A.N., & Arsenin, V.Y. (1977). *Solution of Ill-posed Problems*. Washington: Winston & Sons.

148. Tong, F., Nakayama, K., Moscovitch, M., Weinrib, O., & Kanwisher, N. (2000). Response properties of the human fusiform face area. *Cognitive neuropsychology*, *17*(1-3), 257-280. https://doi.org/10.1080/026432900380607

149. Usher, M., & Cohen, J. D. (1999). Short term memory and selection processes in a frontal-lobemodel. In D. Heinke, G. W. Humphreys & A. Olson (Eds.), *Connectionist models in cognitive neuroscience* (78–91). Berlin, Germany: Springer-Verlag. https://doi.org/10.1007/978-1-4471-0813-9_7

150. Vangkilde, S., Bundesen, C., & Coull, J. T. (2011). Prompt but inefficient: nicotine differentially modulates discrete components of attention. *Psychopharmacology*, *218*(4), 667-680. https://doi.org/10.1007/s00213-011-2361-x

151. Vergauwe, E., & Cowan, N. (2015). Working memory units are all in your head: Factors that influence whether features or objects are the favored units. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(5), 1404-1416. https://doi.org/10.1037/xlm0000108

152. Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*(6984), 748-751. https://doi.org/10.1038/nature02447

153. Vukovic, N., Feurra, M., Shpektor, A., Myachykov, A., & Shtyrov, Y. (2017). Primary motor cortex functionally contributes to language comprehension: an online rTMS study. *Neuropsychologia*, *96*, 222-229. https://doi.org/10.1016/j.neuropsychologia.2017.01.025

154. Wertheimer, M. (1912). Experimentelle studien uber das sehen von bewegung. *Zeitschrift fur Psychologie*, *61*.

155. Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, *131*(1), 48-64. https://doi.org/10.1037/0096-3445.131.1.48

156. Wiegand, I., Töllner, T., Dyrholm, M., Müller, H. J., Bundesen, C., & Finke, K. (2014). Neural correlates of age-related decline and compensation in visual attention capacity. *Neurobiology of Aging*, *35*(9), 2161-2173. https://doi.org/10.1016/j.neurobiolaging.2014.02.023

157. Wiegand, I., Töllner, T., Habekost, T., Dyrholm, M., Müller, H. J., & Finke, K. (2014). Distinct neural markers of TVA-based visual processing speed and short-term storage capacity parameters. *Cerebral Cortex*, *24*(8), 1967-1978. https://doi.org/10.1093/cercor/bht071

158. Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of vision*, *4*(12), 11-11. https://doi.org/10.1167/4.12.11

159. Williams, P., & Simons, D. J. (2000). Detecting changes in novel, complex three-dimensional objects. *Visual Cognition*, *7*(1-3), 297-322. https://doi.org/10.1080/135062800394829

160. Woo, J. Y. T. (2004). *A linguistic introduction to Chinese languages*. Queen Elizabeth School Old Students' Association Education Promotion Organization.

161. Wundt, W. (1899). Zur Kritik tachistosckopischer Versuche. *Philosophische studien*. **15**, 287-317.

162. Xie, W., & Zhang, W. (2017a). Familiarity increases the number of remembered Pokémon in visual short-term memory. *Memory & cognition*, *45*(4), 677-689. https://doi.org/10.3758/s13421-016-0679-7

163. Xie, W., & Zhang, W. (2017b). Familiarity speeds up visual short-term memory consolidation. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(6), 1207-1221. https://doi.org/10.1037/xhp0000355

164. Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233-235. https://doi.org/10.1038/nature06860

165. Zhang, W., & Luck, S. J. (2009). Sudden death and gradual decay in visual working memory. *Psychological science*, *20*(4), 423-428. https://doi.org/10.1111/j.1467-9280.2009.02322.x

166. Zimmer, H. D., Popp, C., Reith, W., & Krick, C. (2012). Gains of item-specific training in visual working memory and their neural correlates. *Brain research*, *1466*, 44-55. https://doi.org/10.1016/j.brainres.2012.05.019

167. Ásgeirsson, Á. G., Nordfang, M., & Sørensen, T. A. (2015). Components of attention in grapheme-color synesthesia: A modeling approach. *PloS one*, *10*(8), e0134456, 1-19. https://doi.org/10.1371/journal.pone.0134456

ISSN (online): 2246-123X
ISBN (online): 978-87-7573-954-7

AALBORG UNIVERSITY PRESS