



On the time required for identification of visual objects

Petersen, Anders

Publication date:
2010

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

[Link back to DTU Orbit](#)

Citation (APA):
Petersen, A. (2010). *On the time required for identification of visual objects*. Technical University of Denmark, DTU Informatics, Building 321. IMM-PHD-2010-224

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

On the time required for identification of visual objects

Anders Petersen

Kongens Lyngby, 2010
IMM-PHD-2010-224

Technical University of Denmark
Informatics and Mathematical Modelling
Building 321, DK-2800 Kongens Lyngby, Denmark
Phone +45 45253351, Fax +45 45882673
reception@imm.dtu.dk
www.imm.dtu.dk

Summary

The starting point for this thesis is a review of Bundesen's theory of visual attention. This theory has been widely accepted as an appropriate model for describing data from an important class of psychological experiments known as whole and partial report. Analysing data from this class of experiments with the help of the theory of visual attention – have proven to be an effective approach to examine cognitive parameters that are essential for a broad range of different patient groups.

The theory of visual attention relies on a psychometric function that describes the ability to identify a stimulus as a function of exposure duration. An important contribution of the thesis is that it investigates whether other psychometric functions than the one originally used with the theory of visual attention could be more appropriate at describing data. The thesis points to two psychometric functions that seem more appropriate. Further the thesis shows that it is possible to incorporate any desired psychometric into the theory of visual attention. Common to the two psychometric functions suggested is that they both have a hazard function that is non-monotonic; a neural argument for this is also presented in the thesis.

For the psychometric function it is further investigated how this depends on stimulus contrast. In this respect, we find that the type of psychometric function is independent of contrast, but that the parameters for the psychometric function vary systematically as a function of contrast.

An analysis of the psychometric function for the individual letters of the alphabet shows that there are significant differences in the parameters of the psychometric function depending on letter identity. Here we should note that in many cases (also for Bundesen's theory of visual attention) it has been customary to average performance over the entire set of stimuli, consisting for instance of the 26 alphabetic letters.

The fact that each letter is perceived in a different way possibly reflects that each letter is represented differently in our brain. This might have to do with a difference in the set of features representing the individual letters. It is possible that some features are processed faster than others and that overlapping features representing more than one letter in the alphabet play a certain role for the tendency to confuse letters. Hopefully it should be possible, with the dataset that we collected, to directly analyse how confusability develops as a certain letter is exposed for increasingly longer time.

An important scientific question is what shapes the psychometric function. It is conceivable that the function reflects both limitations and structure of the physical mechanism underlying perception. For this reason we argue that the alternative psychometric functions that we have suggested are also relevant for models trying to simulate the mechanism leading to perception. The thesis reviews a selection of stochastic models that are well-known candidates when it comes to modelling

mechanisms of perception. These candidates include the Ornstein-Uhlenbeck model and the leaky competing accumulator model.

A further contribution of the thesis is a demonstration that the leaky competing accumulator model (see Usher & Cohen, 1999) is able to explain a perceptual limit that characterises how many objects can in parallel be perceived.

Finally, the thesis suggests five concrete topics for future work. These include as diverse themes as: determination of the visual features representing the individual letters in our brain, neurodynamical modelling of visual perception, investigation of the duration of visual short-term memory as well as psychometric functions and assumptions along with application areas in cognitive diagnostics.

Keywords: *visual identification, exposure duration, non-monotonic hazard function, visual short-term memory, theory of visual attention, cognitive modelling, whole and partial report, psychometric function*

Resumé

Denne afhandling tager udgangspunkt i en kort introduktion til Bundesens teori om visuel opmærksomhed. Denne teori har vundet indpas som en model til beskrivelse af data fra en vigtig klasse af perceptionsforsøg kaldet 'hel og delvis-rapportering'. Disse forsøg har sammen med teorien om visuel opmærksomhed vist sig effektive til at undersøge kognitive parametre, der er væsentlige hos et bredt udsnit af forskellige typer af patientgrupper.

Teorien om visuel opmærksomhed bygger på en såkaldt psykometrisk funktion, der beskriver evnen til at genkende et stimulus som funktion af visningstiden. Et vigtigt bidrag fra denne afhandling er, at den undersøger hvorvidt andre psykometriske funktioner, end den teori om visuel opmærksomhed er født med, måtte være mere velegnede til at beskrive data. Tesen peger på to psykometriske funktioner, der synes mere velegnede, og den viser herudover, hvorledes en vilkårlig psykometrisk funktion i realiteten lader sig indkorporere i teorien om visuel opmærksomhed. Fælles for de foreslåede psykometriske funktioner er, at de har en hazardfunktion, der er ikke-monoton, og der angives i tesen en neural begrundelse for dette.

For den psykometriske funktion er det yderligere undersøgt, hvorledes denne afhænger af stimulus kontrast. I denne sammenhæng har vi fundet, at typen af den optimale psykometriske funktion er uafhængig af kontrast, men at parametrene for funktionen derimod varierer systematisk som funktion af kontrast.

Analyseres den psykometriske funktion for de enkelte bogstaver i alfabetet ses det, at der er signifikante forskelle på den psykometriske funktions parametre i forhold til hvilket bogstav, der er tale om. Dette er værd at bemærke, da det ellers i mange videnskabelige sammenhænge (blandt andet også i forbindelse med Bundesens teori om visuel opmærksomhed) har været sædvane at midle over stimulussættet bestående fx af forskellige bogstaver.

At bogstaverne perciperes forskelligt afspejler muligvis en forskel på de byggesten, der repræsenterer de enkelte bogstaver i vores hjerne. Eksempelvis kunne det tænkes, at nogle byggesten lægges hurtigere end andre, og at overlappende byggesten, der indgår i mere end et bogstav i alfabetet, spiller en rolle for tilbøjeligheden til at forvirre de enkelte bogstaver med hinanden. Med det opsamlede datasæt er det for første gang blevet muligt direkte at analysere, hvorledes forvirringen mellem bogstaver udvikler sig efterhånden som et bogstav vises i længere og længere tid.

Et vigtigt videnskabeligt spørgsmål er, hvad der former den psykometriske funktion. Det er oplagt, at funktionen afspejler både begrænsninger og strukturen af den fysiske mekanisme, der muliggør visuel perception. Derfor har vores forslag til alternative psykometriske funktioner også relevans for modeller, der forsøger at simulere den mekanisme, der leder til perception. I tesen gennemgås en række stokastiske modeller der er velkendte kandidater, når det drejer sig om at modellere perceptionsmekanismer.

Til de udvalgte kandidater hører Ornstein-Uhlenbeck modellen og den lækkende konkurrerende akkumulator model.

Et yderligere bidrag i tesen er en demonstration af, at en lækkende konkurrerende akkumulator model (se Usher & Cohen, 1999) kan forklare den perceptuelle begrænsning der ligger i, hvor mange objekter der parallelt lader sig opfatte.

Endeligt foreslår tesen fem konkrete fremtidig arbejdsområder. Disse omfatter så forskelligartede emner som bestemmelse af de byggesten der repræsenterer de enkelte bogstaver i vores hjerne, neurodynamisk modellering af visuel perception, undersøgelse af varigheden af visuel korttidshukommelse såvel som psykometriske funktioner og antagelser samt anvendelsesområder i kognitiv diagnostik.

Preface

This thesis was prepared at the Technical University of Denmark (DTU), Department of Informatics and Mathematical Modelling (IMM) in the period from September 2006 to November 2009. The thesis serves as partial fulfilment of the requirements for acquiring the Ph.D. degree in engineering. The project has been carried out in collaboration with the Center for Computational Cognitive Modeling (CfCCM) and the Center for Visual Cognition (CVC), that are both located at the Department of Psychology, University of Copenhagen (CU).

The thesis was generously funded by a full stipend from the Oticon Foundation.

Supervisors on the project were: Associate Professor Tobias Søren Andersen (DTU), Professor Lars Kai Hansen (DTU) and Associate Professor Søren Kyllingsbæk (DTU and CU).

The thesis deals with mathematical modelling of visual perception. It uses tools such as statistics, probability theory and optimization to model original behavioural data.

The thesis consists of a summary report and two research papers (A. Petersen & Andersen, 2009; A. Petersen, Kyllingsbæk, & Hansen, 2009) that were prepared during the Ph.D. study.

Kgs. Lyngby, November 2009



Anders Petersen

Acknowledgements

This is the time to thank everyone that made this thesis possible, so first of all let me thank my three supervisors.

Søren Kyllingsbæk I thank for his perfectly developed logical way of thinking, his ability to empathize, his open-mindedness and his willingness to introduce me to the interesting research carried out at CVC.

Lars Kai Hansen I thank for his vibrant personality, his astounding omniscience, his never-failing sanguinity as well as his incomparable mathematical clear-sightedness.

Finally, but not the least, I thank my main supervisor Tobias Søren Andersen for introducing to me psychophysics in practise, for recognising the importance of my seemingly over-elaborate experimental efforts, for his meticulous but sound critique of my writing, for his eagerness to discuss scientific matters, for his inspiration for generalizing and structuring my results and for his many caring attempts to develop me into a better researcher. Also I thank Tobias for sharing with me enjoyable anecdotes and colourful impressions from his many experiences from abroad.

Further, I want to thank Hitomi Shibuya, Claus Bundesen & Lisbeth Harms for access to the experimental data in (Shibuya & Bundesen, 1988; Bundesen & Harms, 1999), for warm-heartedness very hard to miss, and for helpful comments on early versions of my work. As I found so much inspiration in your interesting work, I hope that you in particular will find my work to be worthwhile.

I would also like to thank my colleagues at IMM and CVC; in part for good company in Denmark and abroad, and in part for discussing with me various mathematically useful matters. Especially I would like to thank Mikkel Nørgaard Schmidt for interesting discussions on Bayesian statistics, Gaussian processes etc.

Finally I warmly thank my friends and especially my family to whom I owe everything. I am sure that the many years that I spent studying and later preparing my Ph.D. may have caused you to suffer some deprivation. I hope however, that you will forgive me this. Also I wish that you will take pride in the fact that your deprivation enabled my thesis. To the best of my ability the thesis improves current scientific standards - as pompous as it sounds - hopefully for the common-good of mankind.

Thank you all..!

Nomenclature

Abbreviations

Abbreviation	Explanation
AD	Alzheimer's Disease
AIC	Akaike Information Criterion
BIC	Bayes Information Criterion
CfCCM	Center for Computational Cognitive Modeling
CVC	Center for Visual Cognition
CU	University of Copenhagen
DTU	Technical University of Denmark
FIRM	Fixed-capacity Independent Race Model
HD	Huntington's Disease
IMM	Informatics and Mathematical Modelling
LCA	Leaky Competing Accumulator
MCI	Mild Cognitive Impairment
NTVA	Neural Theory of Visual Attention
TVA	Theory of Visual Attention
VSTM	Visual Short-Term Memory
OU	Ornstein-Uhlenbeck

Symbols

Symbol	Explanation
A	Activation
b	Barrier
C	Total processing capacity
D	Number of distractors
dw	White noise contribution
f	Probability density for target encoding
F	I) Psychometric function after correction for guessing and lapsing II) Distribution function target encoding II) Activation function
g	Probability density for distractor encoding
G	I) Distribution function target encoding II) Poisson spike train
j	The score, i.e. the number of reported targets
k	A psychometric parameter
K	Storage capacity of the VSTM
S	Visual field
t	Time
T	Number of targets
v	Element processing rate
V	A psychometric parameter
X	Stochastic variable
Y	Stochastic variable
α	I) Ratio of processing resources II) Self-excitation
β	Lateral inhibition
γ	I) Guessing rate II) Amplitude scaling of Poisson spike train
θ	Set of psychometric parameters
λ	Lapsing rate
λ_L	Leakage term
μ	I) Drift term / II) a psychometric parameter
σ	I) Noise amplitude / II) a psychometric parameter
τ	A psychometric parameter
Ψ	Psychometric function before correction for guessing an lapsing

Note: Roman numbers indicate symbols that can represent multiple entities.

Contents

Summary.....	i
Resumé	iii
Preface	v
Acknowledgements	vii
Nomenclature.....	ix
Abbreviations.....	ix
Symbols.....	x
Contents.....	xi
1 Introduction.....	1
1.1 Understanding the human visual system	1
1.2 Visual Short-Term Memory	1
1.3 Vision experiments.....	2
1.4 Visual displays for diagnosis.....	2
1.5 The utility of single, whole and partial report related to diagnosis of AD.....	3
1.6 Thesis overview	4
PART I: THEORY	7
2 The Theory of Visual Attention.....	9
2.1 Encoding: sharing of processing resources between elements.....	9
2.2 Storage: the fixed-capacity independent race model.....	10
3 Single-object report.....	13
3.1 A notion on identification and the first-passage-time distribution.....	13
3.2 Alternative psychometric functions.....	14
4 The leaky competing accumulator model.....	19
PART II: METHODS AND RESULTS	23
5 Research question	25
6 The research themes.....	27
6.1 Single-letter report: the psychometric function.....	27
6.2 Whole and partial report: the generalized FIRM equations	30
6.3 The varying storage capacity of the visual short-term memory	33
6.4 Individual identities: is averaging okay?.....	34
6.5 Confusion: does it develop with exposure duration?	37

6.6	Contrast: does it affect the form of the psychometric function?	38
6.7	Guessing rates: what do they reveal?	39
6.8	Summary on our principal research question - the psychometric function	42
PART III: CONCLUSION		45
7	Conclusion	47
7.1	Summary of contributions	50
7.2	Directions for future work	51
Appendix A	Paper for ICCM2009	55
Appendix B	Paper submitted	63
References	119

1 Introduction

1.1 Understanding the human visual system

Most people would probably agree that vision is one of the most important senses that humans have; the fact that it is the sense that occupies the largest part of the brain seems to support this point of view. There are many reasons why one would want to model and understand the functioning of the human visual system. Important reasons include scientific curiosity, understanding disorders and pathologies linked to the visual system and inspiration for synthesis of artificial vision systems. In this thesis we focus on the development of models that could potentially lead to better diagnostic instruments for patient monitoring. The models that we develop are frequently aimed at assessing cognitive disorders, and so we would be very satisfied if the results of this thesis can improve on these very useful models. If our models also serve to provoke scientific curiosity, for instance serving as inspiration for future research, we would of course be happy for this also.

1.2 Visual Short-Term Memory

In many visual tasks performance rely on limitations of a subjects visual short-term memory. The visual short-term memory is part of the human working memory system, and it plays a key role for visual identification.

It is not enough for identification of an object that it causes sensory input, leading to what is known as an iconic image; if representation of the visual objects is not upheld by either continued visual presentation or executive control from higher cortical areas, memories in iconic memory will simply decay away inside a temporal span of less than a second (Sperling, 1960). Iconic memories are fragile and decay rapidly, whereas visual short-term memories can be robust to subsequent stimuli and last for a period of many seconds (Phillips & A. D. Baddeley, 1971).

As the concept of Visual Short-Term Memory (VSTM) is crucial in our understanding of how we identify objects, it might seem important how we define it. We consider VSTM as a mechanism temporarily capable of holding on to a limited number of selected visual objects. The object that we consider is either a digit or a letter entity, more generally objects could also be defined as ‘chunks’ of visual information (Miller, 1956), hence in principle a word or a number consisting of more than one digit could also count as a visual object. When an object enters VSTM, this could be interpreted as a population of neurons (representing the object) having its activation sustained by being incorporated into a type of feedback loop (Bundesen, Habekost, & Kyllingsbæk, 2005, p. 302).

According to the Theory of Visual Attention (TVA; Bundesen, 1990) the visual short-term memory is limited by both processing capacity as well as storage capacity. Already in the late 19th century a surprising limit in how many objects that can be perceived at the same time was demonstrated. It seems that only about 4 objects can be held in the VSTM at the same time (Cattell, 1886; Cowan, 2000). This finding is independent of the number of objects visually presented (Sperling, 1960).

Evidence further exist that the “magical number” of 3-to-4 objects is largely independent of how many features that are encoded for each object. This means that the complexity of the visual object does not hold an influence on the memorial capacity of the VSTM see (Luck & Vogel, 1997), but see also (Alvarez & Cavanagh, 2004). This later empirical finding is consistent with Bundesen’s Theory of Visual Attention (Bundenen, 1990) which assumes that VSTM is limited by the number of visual objects rather than by the number of features that constitute the object.

1.3 Vision experiments

Many different types of psychological experiments have been developed through the last few centuries, and again many of these paradigms are aimed at assessing various aspects of human vision. As a visual scene typically contains several objects, one or more which are of special importance for us to identify and some that are not, it seems vital for the visual system how well it behaves in such tasks. For this reason we choose to narrow down our focus to a class of experiments known as whole and partial report. Whole and partial report aim directly at quantifying the accuracy of visual object identification in single- and multi-object displays, here accuracy is assumed to be heavily dependent on the subject’s visual short-term memory limitations.

1.4 Visual displays for diagnosis

There is a long history for applying visual displays (e.g. consisting of letters) to diagnose disorders affecting the visual system. These visual disorders may originate anywhere from the peripheral system, through the neuro-sensory pathway even to the cortical regions engaged in vision. We shall now give examples that visual displays can be used to assess disorders originating virtually at any level in the visual system:

Our first example is the Snellen letter chart (Snellen, 1862) which is in fact very well known. Even today a Snellen letter chart is situated on a wall in most medical clinics; the Snellen chart is known as a convenient tool for diagnosing short and long sightedness, which most often originates in a peripheral, eye disorder.

Our second example shows that visual letter displays are also useful for diagnosing disorders originating in the visual pathway. This was demonstrated when Jannik Petersen Bjerrum used multiple Snellen type letters printed in various contrast levels to diagnose patients suffering from visual deficits in various regions of the visual field (Bjerrum, 1882, 1889; Dreyer, Edmund, & Møller, 1992). Bjerrum discovered the arcuate scotoma in glaucoma, and he later earned the reputation as the founder of

campimetry, due to his efforts to develop methods for measuring a subject's field of vision.

Our third and last example shows that visual letter displays are furthermore useful for diagnosing cognitive disorders that could originate in visual cortical regions. In paradigms such as whole and partial report the subject has to report multiple simultaneously presented letters that are displayed as a function of exposure duration. In fact, it has been shown that both whole and partial report provide correlates of the patient's state under a number of cognitive disorders (Habekost & Starrfelt, 2008), these disorders include Alzheimer's disease (Bublak, Redel, & Finke, 2006).

1.5 The utility of single, whole and partial report related to diagnosis of AD

This section gives an example of how single, whole and partial report can possibly be useful for diagnostic purposes, exemplified by diagnosis of Alzheimer's disease (AD). Today 5.3 million people are living with AD in USA; every 70 seconds a person develops AD; and further AD and dementia triple the cost for healthcare for persons of age 65 years and older (Alzheimer's Association, 2009). For the sake of early diagnosis as well for allowing for patient monitoring when the patient is undergoing medical treatment, it is important to have the best possible set of indicators that can reliably predict and assess the state of AD. Recently various types of bio-markers have been shown to predict some cases of AD (Mattsson et al., 2009; R. C. Petersen & Trojanowski, 2009), but there is still plenty of room for other types of test procedures, especially some that are both non-invasive and can detect the disease at a very early point. Cognitive test batteries used for this purpose include the Mini-Mental State Examination and the Alzheimer's Disease Assessment Scale (Mattes, 1997; M. F. Folstein, S. E. Folstein, & McHugh, 1975; Wouters, van Gool, Schmand, & Lindeboom, 2008); these are both popular and fairly quick tools for judging cognitive function. In order to be able to improve on such test batteries, and because AD affects more and more people, we agree with Bublak et al. (2006, 2009) and Habekost & Starrfelt (2008) that it is interesting that human performance in whole report displays (displays with a number of simultaneous targets) have been shown to predict AD and even a similar but milder state known as mild cognitive impairment (MCI).

In whole-report displays, predictors of AD and MCI were the minimum exposure time (the offset) needed for identification of an object and further for AD a leftward bias of spatial attention (Bublak et al., 2006; Bublak et al., 2009). Both of these predictors also served to predict Huntington's disease (HD). Related to partial report (which also includes distractors in the visual display) there is evidence that the ability to filter out irrelevant elements in the visual field is also an indicator of AD (A. D. Baddeley, H. A. Baddeley, Bucks, & Wilcock, 2001). Further, for single-object identification, a large scale study involving hundreds of subjects showed that AD and MCI was quite strongly predicted by the median time for identifying objects. In a comparison of a battery of 16 different cognitive tasks (Mendola, Cronin-Golomb, Corkin, & Growdon, 1995) visual identification of letters (followed by a so-called backward pattern mask, as is standard in the whole and partial report paradigm) was one of the many experimental procedures

that were tested. For this task, the exposure duration at the 50 % correct threshold was shown to yield the best predictive power of AD compared to the measures calculated from any other of the cognitive tasks that were tested. A late 50 % threshold was prevalent in 58 % of AD subjects. Whole and partial report procedures can be applied to capture all of the mentioned indicators of AD and MCI provided a model such as the theory of visual attention is used to analyse the data.

The fact that single-letter report showed out to be a good predictor of AD (Mendola et al., 1995; Cronin-Golomb, Corkin, & Growdon, 1995; Cronin-Golomb & Hof, 2004; Cronin-Golomb, Gilmore, Nearing, Morrison, & Laudate, 2007) makes this experiment particularly interesting to study in further detail. We speculate that imposing an appropriate model (a psychometric function), when fitting single-letter data, might provide a more efficient and even more accurate predictor in terms of AD. Also we think that it might be useful to combine the parameters from the psychometric function, such as the threshold, the slope and perhaps an offset to build the most effective estimator of the cognitive deficit, as this is evidenced in for instance AD.

1.6 Thesis overview

The thesis is organized in three parts, which we shall now present to the reader.

PART I: THEORY

In Chapter 2 we introduce central aspects of Bundesen's theory of visual attention. Particularly, in Section 2.1 we see how the theory deals with limits in processing capacity and in Section 2.2 we see how it deals with limits in storage capacity.

In Chapter 3 we first present (in Section 3.1) a few notions on visual identification with the aim of linking the first-passage-time distribution to the psychometric function in single-object report. In Section 3.2 we provide a collection of psychometric functions that we aim to evaluate with respect to their ability to model character identification as a function of exposure duration.

In Chapter 4 we present the leaky competing accumulator model, which is a type of diffusion model that can be used to model a cognitive process such as the identification task. We shall later use the model to provide a neural network demonstration how visual short-term memory storage capacity can vary from trial to trial.

PART II: METHODS AND RESULTS

In Chapter 5 we formulate our research question which is motivated by the introduction of visual identification that we gave in Chapter 1 and the theories and models related to visual identification that we presented in Chapters 2 to 4. After having stated our research question we present seven related research themes to which we believe this thesis will contribute.

In Chapter 6 we present the various investigations that we conducted to answer our research question. Our research is treated theme by theme; and each of the seven themes is dedicated its own section, which briefly explains what was the method used to treat the research question and what was the result from this investigation. Further details of methods and results can be found in the two research papers (A. Petersen et al., 2009) and (A. Petersen & Andersen, 2009) that were prepared during the course of the Ph.D.

PART III: CONCLUSION

In Chapter 7 we discuss implications for each of the results obtained in 5. We conclude on our results, relate them if possible, and discuss how we tried to anticipate possible confounds. Finally we point out general tendencies seen across various datasets and discuss implications of such convergence. Also as part of the conclusion we point out the most important contributions of our research, and conclude on whether our work has addressed the research question. Finally, we suggest some topics for future work hoping that as many as possible of these topics will be treated in future studies.

Finally, we also include Appendix A which contains our conference paper: (A. Petersen et al., 2009) as well as Appendix B that contains our submitted manuscript: (A. Petersen & Andersen, 2009).

References can be found in the last pages of the thesis.

PART I: THEORY

2 The Theory of Visual Attention

This chapter, which is to a large extent an excerpt from Appendix B in (A. Petersen & Andersen, 2009) presents how the Theory of Visual Attention (TVA; Bundesen, 1990) can be used to provide an account of performance in single, whole and partial report. The tradition in TVA is to assume that guessing and lapsing can be ignored (Bundesen, 1990; Bundesen & Harms, 1999; Shibuya & Bundesen, 1988) this simplifies matters considerably especially in whole and partial report (Shibuya & Bundesen, 1988).

In TVA any target or distractor is denoted as an *element*, and further it is assumed that a subject only correctly identifies the elements that are stored in visual short-term memory. Further, subjects only obtain a chance to *store* an element if the element is *encoded*.

2.1 Encoding: sharing of processing resources between elements

The hazard rate that a particular element, i , is encoded into VSTM is proportional to how large a portion of processing resources the element receives. Any element in the visual field S receives a certain portion v_i of the total processing capacity C , which is assumed to be invariant with respect to the number of elements in the display:

$$C = \sum_{i \in S} v_i$$

It serves as a simplification to assume homogeneity of the visual display (Shibuya & Bundesen, 1988). This appears a reasonable assumption as long as all elements have the same size, the same contrast, the same eccentricity etc. The homogeneity assumption means that all targets receive the same amount of processing resources denoted v_t . In the same way all distractors receive the same amount v_d of processing resources which is proportionally smaller than the amount that targets receive. The ratio of processing resources α is defined as:

$$\alpha = \frac{v_d}{v_t}$$

Let us assume a homogenous display that contains T targets and D distractors. The processing resources v_t of any target is then given by:

$$C = \sum_{i \in S} v_i \Rightarrow C = Tv_t + Dv_d = Tv_t + \alpha Dv_t = (T + \alpha D)v_t \Leftrightarrow v_t = \frac{C}{T + \alpha D}$$

2.2 Storage: the fixed-capacity independent race model

TVA describes the factors that determine the probability that any given target in a multi-element visual display is encoded into visual short-term memory, however as VSTM typically has only about 3-4 storage places, TVA assumes that not all elements that become encoded are actually stored. Bundesen (1990) assumes that the occupation of places occurs through a so-called race, that is, any newly encoded element will lead to immediate occupation of one storage place if and only if there is still any storage place left in the VSTM.

Let us define that f and F and are the probability density and the distribution function for target encoding respectively. Similarly we also define that g and G are the probability density and the distribution function for distractor encoding.

According to the fixed-capacity independent race model (Shibuya & Bundesen, 1988) the probability of a score of j (targets reported) from a display containing T targets and D distractors exposed for t_e seconds can be written as:

$$P(j; T, D, t_e) = P_1 + P_2 + P_3$$

where P_1 is the probability that the score equals j and the total number of elements (targets and distractors) entering VSTM is *less* than K . The number of distractors entering VSTM is denoted by m and $m \leq \min(D, K-j-1)$. If $j=K$, $P_1=0$; otherwise:

$$P_1 = \binom{T}{j} [F(t)]^j [1-F(t)]^{T-j} \cdot \sum_{m=0}^{\min(D, K-j-1)} \binom{D}{m} [G(t)]^m [1-G(t)]^{D-m}$$

and P_2 is the probability that the score equals j and the total number of elements *equals* K and the K^{th} element entering the VSTM is a *target*. The number of distractors entering VSTM denoted by m is always $K-j$. If $j=0$, or $j < K-D$, $P_2=0$; otherwise:

$$P_2 = \int_0^{\tau} \binom{T-1}{j-1} [F(t)]^{j-1} [1-F(t)]^{T-j} \binom{D}{m} [G(t)]^m [1-G(t)]^{D-m} \binom{T}{1} f(t) dt$$

and P_3 is the probability that the score equals j and the total number of elements *equals* K and the K^{th} element entering the VSTM is a *distractor*. The number of distractors entering VSTM denoted by m is always $K-j$. If $j=K$, or $j < K-D$, $P_3=0$; otherwise:

$$P_3 = \int_0^{\tau} \binom{T}{j} [F(t)]^j [1-F(t)]^{T-j} \binom{D-1}{m-1} [G(t)]^{m-1} [1-G(t)]^{D-m} \binom{D}{1} g(t) dt$$

Shibuya and Bundesen (1988) derived explicit score probabilities under the assumption that encoding proceeds as a homogenous Poisson process. This corresponds to assuming (ignoring the temporal offset) that the hazard rates are constant over time. An original contribution from our work is a generalized model that allows the hazard rates to be time-varying, although the generalized model still assume that the hazard rates (for the

different elements presented) are mutually proportional functions of time (cf. Bundesen, 1990, 1993, 1998). Explicit formulas for our so-called generalized Fixed-capacity Independent Race Model (FIRM) as well as an explanation of how we relaxed the assumption on the hazard rates are available in Appendix B in (A. Petersen & Andersen, 2009).

3 Single-object report

In Section 3.1 we shall give an interpretation of visual identification and introduce a few practical assumptions about it, including what is required for identifying an object. In Section 3.2, which is to a large extent an excerpt from the introduction in (A. Petersen & Andersen, 2009) we present a selection of psychometric functions that we use later on when we shall model several datasets containing data from single, whole and partial report experiments.

3.1 A notion on identification and the first-passage-time distribution

Identification as such can be regarded as a cognitive state (or simply a choice) that is the result of some neuro-sensory judgement process. For instance the subject could have to judge if a certain stimulus category is present or not. This judgement process is dependent on how much information about the stimulus that has been accrued. Different models of perception vary considerable in both how information builds up and what is the decision rule that is used for the judgement. For simplicity we shall also assume that the sensory input, that is, the visual field do not change during the time course in which stimuli are presented. The only thing that does change is how long time the stimulus is presented.

The assumption that an *information threshold* must be crossed for inference of the stimulus identity is an important aspect of many types of cognitive models, such as response time models (Smith, 2000). We also assume that if and only if an information threshold is crossed the stimulus identity is inferred. For simplicity we shall assume a high-threshold for identification. The high-threshold assumption implies that either the subject is able to infer the correct identity of the target or else the subject guesses randomly. If we denote the probability of inferring the correct identity of the stimulus as $P(id)$ and the probability of guessing the correct identity of the stimulus as γ , then the probability that the subject will report the correct identity of the stimulus, $P(hit)$, is:

$$P(hit) = P(id) + \gamma(1 - P(id)) = \gamma + (1 - \gamma)P(id)$$

The high-threshold assumption implies that the subject never infers a false identity of the target, and therefore false categorizations can therefore solely be explained as the result of random guessing. Though we decided to follow the popular high-threshold assumption; it might be interesting to explore a low-threshold assumption as well, as it is sometimes produces slightly more realistic predictions, see (Palmer, Verghese, & Pavel, 2000). There are several low-threshold models to choose from (Palmer et al., 2000) however an important subset of these is based on signal detection theory - for instance the ideal observer model (Green & Swets, 1966; Geisler, 2003). What is common to many low-threshold models is that they assume noisy representations of the stimuli as well as a decision making process that gradually integrates information about

the stimulus categories and incorporates various kinds of attentional biases. Let us emphasize that here we do not deal with low-threshold theories as we follow the simpler high-threshold assumption as we have already explained above.

The *first-passage-time distribution* is important to keep in mind when modelling behavioural types of experiments such as the identification task. The first-passage-time is the time, t_b at which a stochastic process, $A(t)$ reaches or crosses a certain barrier (threshold) level, b for the first time.

$$t_b = \inf \{t \geq 0 ; A(t) \geq b\}$$

In relation to a psychological experiment such as the identification task we are most often not able to observe exactly how information builds up, or even how this depends on various stimulus features or other experimental settings. Given our above mentioned assumption what we can deduce from our observations however is whether the information that has built-up has exceeded the information threshold for a particular trial. Mathematically the probability distribution of crossing a certain information threshold corresponds to the cumulative first-passage-time distribution.

From a modelling point of view single-object report appears quite similar to the classical response time problem as far as single-object report is also temporally dependent paradigm. Therefore also for single-object report the distribution of processing times can be interpreted the first-passage-time distribution of an underlying stochastic process of visual identification (Smith, 2000).

3.2 Alternative psychometric functions

When considering the single-object report experiment the proportion correct is directly characterized by the so-called psychometric function. In a modeling perspective the shape of the psychometric function depends on the type of stochastic accumulation model that is assumed. For many types of stochastic accumulation models (and also for TVA) it is often assumed that one crossing of the activation threshold is sufficient for inferring the category (Smith, 2000; Whitmore, 1986; Aalen, Borgan, & Gjessing, 2008; Bundesen, 1990); for single-object report this means that the first-passage-time distribution equals the psychometric function (cf. Section 3.1). A reason why we think the psychometric function is important from a cognitive modeling perspective is that it allows us to narrow down the field of cognitive model candidates to be considered.

A psychometric function $\psi(t; \theta, \gamma, \lambda)$ quantifies the probability of a correct report as a function of some stimulus attribute t , which in our case is exposure duration. It is characterized by a number of parameters that include the parameter set θ of the function F as well as two additional parameters, γ and λ , that denote the guessing and lapsing probabilities, respectively. We define the guessing probabilities as the fraction of times an un-informed observer presses (intentionally or accidentally) each of the keys included in the response set. The lapsing probability we define as the relative fraction of accidental key presses, averaged over all keys in the response set. The psychometric function ψ , which includes correction for guessing and lapsing, can be written as

$$\psi(t; \theta, \gamma, \lambda) = F(t; \theta) \cdot (1 - \lambda) + (1 - F(t; \theta)) \cdot \gamma = \gamma + (1 - \gamma - \lambda) \cdot F(t; \theta)$$

What we shall generally speak of as the *psychometric function* is the function F , i.e. the psychometric function *after* correction for guessing and lapsing (Treutwein & Strasburger, 1999; Wichmann & Hill, 2001).

The exponential distribution with a temporal offset included was used as the psychometric function in TVA (Bundesen, 1990). Ignoring the temporal offset, this psychometric function is an implication of assuming that encoding into VSTM can be considered events from a *homogenous* Poisson process, for which the waiting time is well known to be exponentially distributed. The exponential distribution has the parameter set $\theta = \{\nu, \mu\}$, where $\nu > 0$ is the rate and $\mu > 0$ is the temporal offset of the Poisson process, and the distribution is defined by:

$$F(t; \theta) = 1 - e^{-\nu(t-\mu)} \quad \text{for } t \geq \mu \quad \text{and} \quad F(t; \theta) = 0 \quad \text{for } t < \mu$$

From a psychophysical perspective, it is strange that there should be a fixed temporal offset before encoding can take place, and that after this point the encoding rate is constant; instead, we find it more plausible that the encoding rate rises as a smooth function of exposure duration. The gamma, the Weibull and the ex-Gaussian distributions all represent generalisations of the exponential distribution, and all of these are smooth functions.

The Weibull distribution has the parameter set $\theta = \{\mu, \sigma, k\}$, where $\mu, \sigma, k > 0$. It reduces to the exponential distribution when the shape parameter $k = 1$. The Weibull distribution has previously been used for modelling the psychometric function of visual contrast detection, visual discrimination as well as visual identification when these were investigated as a function of stimulus contrast (Pelli, 1985, 1987; Pelli, Burns, Farell, & Moore-Page, 2006). When the Weibull distribution has $\mu > 0$ it includes an offset. The three-parameter Weibull distribution function is defined as:

$$F(t; \theta) = 1 - e^{-\left(\frac{t-\mu}{\sigma}\right)^k} \quad \text{for } t \geq \mu \quad \text{and} \quad F(t; \theta) = 0 \quad \text{for } t < \mu$$

The gamma distribution has the parameter set $\theta = \{\mu, \sigma, k\}$, where $\mu, \sigma, k > 0$. The gamma distribution corresponds to the waiting-time distribution, when waiting for k independent, and identically distributed events, that each has an exponentially distributed waiting time-distribution and it thus reduces to the exponential distribution when $k = 1$. If correct identification depends on the firing of several independent neural units firing as Poisson processes, then the gamma distribution could describe the psychometric function of identification as a function of stimulus duration. Based on a similar argument the gamma distribution has been fitted to response time distributions (Van Breukelen, 1995; Luce, 1991). Noting that Γ is the complete gamma function, and γ is the lower incomplete gamma function, the three-parameter gamma distribution function is defined as:

$$F(t; \theta) = \frac{\gamma\left(k, \frac{t - \mu}{\sigma}\right)}{\Gamma(k)} \quad \text{for } t \geq \mu \quad \text{and} \quad F(t; \theta) = 0 \quad \text{for } t < \mu$$

The Ex-Gaussian distribution has the parameter set $\theta = \{\mu, \sigma, \tau\}$, where $\mu, \sigma, \tau > 0$. The ex-Gaussian approaches the exponential distribution in the limit, to be exact when $\mu = 0$ and $\sigma \rightarrow 0$. The Ex-Gaussian distribution characterises the sum of an exponential distributed variable and a Gaussian distributed variable. Thus, if Gaussian noise is added to the temporal offset, μ , in the exponential distribution the waiting times for perceptual processing would be distributed according to the ex-Gaussian distribution. The ex-Gaussian has been used for modelling reaction-time data (Luce, 1991). Noting that Φ is the Gaussian distribution function, the ex-Gaussian distribution function is defined as:

$$F(t; \theta) = \Phi\left(\frac{t - \mu}{\sigma}\right) - \Phi\left(\frac{t - \mu - \sigma^2 / \tau}{\sigma}\right) \cdot \exp\left(-\frac{t}{\tau} + \frac{\mu}{\tau} + \frac{\sigma^2}{2\tau^2}\right)$$

Also, an important question is what causes the shape of a psychometric function? Clearly the shape must reflect the construct and limitations of the physical mechanism underlying perception, i.e. it must reflect the neural activity in the task relevant areas of the brain. It has previously been demonstrated that individual sensory neurons show response functions (firing rate vs. stimulus intensity) that closely resemble the psychometric function seen in detection tasks, such as the logistic distribution (P. Lansky, O. Pokora, & J. P. Rospars, 2007). In NTVA, which is a neural interpretation of TVA (Bundesen et al., 2005), it is assumed that the rate at which stimuli are perceptually processed is proportional to neural firing rates in the visual cortex. Mathematically, the processing rate is the *hazard* rate. Further descriptions of the concept of hazard rate can be found in (Luce, 1991; Van Zandt, 2002; Aalen & Gjessing, 2001). The processing rate is explicit in the exponential function where it equals the parameter, ν . For other psychometric functions the hazard rate is generally not explicit but can easily be derived (see Appendix B).

Exposing cats and monkeys to transient stationary gratings with a duration of 200 ms, (Albrecht, Geisler, Frazor, & Crane, 2002) mapped out the instantaneous firing rates of responsive neurons in the visual striate cortex. The typical temporal profile of the firing rates is similar to the profile that (Bundesen & Habekost, 2008, p. 116) expect follows the abrupt onset of a stimulus: ‘*When a stimulus appears abruptly (a kind of successive contrast), firing rates of typical neurons responding to the stimulus first increase rapidly, then reach a maximum, and finally decline and approach a somewhat lower, steady state level*’. Therefore, it is likely that the psychometric function for object identification as a function of exposure duration has a non-monotonic hazard rate. Accordingly, a preliminary report (Shibuya, 1994) described the hazard function in a 2-AFC discrimination task, in which exposure duration was varied, as having a non-monotonic hazard rate. However, all of the functions described above have monotonic hazard rates. Therefore, we find it worthwhile to consider also two psychometric functions that have similar temporal profiles, i.e. non-monotonic hazard functions.

The log-logistic is such a distribution, since with appropriately chosen parameters; it has a uni-modal, and hence non-monotonic hazard function. The Log-logistic (or Fisk) distribution is the probability distribution of a random variable whose logarithm has a logistic distribution. It has been used for modelling various kinds of diffusion processes (Brüederl & Diekmann, 1995; Diekmann, 1992). Also it has been used for modelling proportion correct in single-digit identification as a function of contrast (Strasburger, 2001). The Log-logistic distribution has the parameter set $\theta = \{\mu, \sigma\}$, where $\mu, \sigma > 0$. Noting that the parameter σ determines the steepness, and μ is the median survival time, the log-logistic distribution function is defined as:

$$F(t; \theta) = \frac{1}{1 + \left(\frac{t}{\mu}\right)^{-\sigma}}$$

The squared-logistic distribution is another distribution with a non-monotonic hazard function. We describe the squared-logistic because we found it to be a simple function which has a hazard function that closely resembles the instantaneous firing rate of single neurons in the visual cortex like those depicted in (Albrecht et al., 2002). Compared to the hazard function of the log-logistic distribution the hazard function of the squared-logistic distribution seems to drop off faster after the peak, and furthermore the hazard approaches the quasi-stationary level V rather than continuing to drop off as $t \rightarrow \infty$. The squared-logistic has the parameter set $\theta = \{V, \mu, \sigma\}$, where $V, \mu, \sigma > 0$. We define the squared-logistic distribution function as:

$$F(t; \theta) = 1 - e^{-\frac{V \cdot t}{\left(1 + e^{\left(\frac{t-\mu}{\sigma}\right)^2}\right)^2}}$$

We named the distribution the squared-logistic because, the shape of the mean cumulative hazard function in the interval between 0 and t as a function of time has the shape of a logistic distribution function squared. This can be seen by dividing the negative exponent (the cumulative hazard function) of the distribution by the size t of the temporal interval. Note that V scales the hazard rate and that it is straightforward to derive the probability density function if needed.

4 The leaky competing accumulator model

In this chapter we show how the leaky competing accumulator model can be developed from more basic type of diffusion processes. As we shall later see in Section 6.3 a leaky competing accumulator model can be used to model a limited visual short-term memory storage capacity. Thereby the leaky competing accumulator model represents an alternative approach to model the storage capacity limitation in connection with TVA.

A *diffusion process* can be described as the solution of a stochastic differential equation (Karlin & Taylor, 1975). One of the simplest and most well-known diffusion processes is the Wiener process. (In physics the Wiener process is sometimes known as *standard Brownian motion*).

$$dA = \sigma \cdot dw$$

For the Wiener process we note that the solution as a function of time, $A(t)$ results as the integration of independent white noise contributions dw with amplitude σ . If a drift term, μ is included we obtain a drift-diffusion process (also known as a Wiener process with drift, a Brownian motion with drift or simply a Brownian motion).

$$dA = \mu \cdot dt + \sigma \cdot dw$$

The first-passage-time distribution for crossing a positive threshold assuming a Brownian motion is the inverse-Gaussian distribution (Singpurwalla, 1995; Lee & Whitmore, 2006).

A generalization of the drift-diffusion process, and also an example of a diffusion process is the Ornstein-Uhlenbeck (OU) process (Ricciardi & Sacerdote, 1979). The OU-process has been widely used to model various biological processes such as neuronal responses.(Ditlevsen, 2007; Petr Lansky & Ditlevsen, 2008; Petr Lansky, Ondrej Pokora, & Jean-Pierre Rospars, 2008; Aalen & Gjessing, 2004).

$$dA = [\mu - \lambda_L \cdot A] \cdot dt + \sigma \cdot dw$$

The OU includes three parameters governing the mean reverting rate, λ_L , the long-term mean, μ and the volatility, σ of the process, respectively. When modelling neuronal responses the spike-to-spike interval is modelled as the first-passage-time distribution. Though the OU-process has been intensively investigated and an analytic solution to the first-passage-time distribution was given in (Ricciardi & Sato, 1988) still no closed-form expression exists for the first-passage-time distribution (Nobile, Ricciardo & Sacerdote, 1985). Other models, such as the Feller process (Ditlevsen, 2007) can also be used for modelling neural firing rates, however as for the OU process few of these have a closed-form solution for the first-passage-time distribution, for an exception however

see (Crescenzo & Martinucci, 2007). It is well-known however that the shape of the first-passage-time distribution depends heavily on the parameters of the OU-process; see e.g. (Aalen & Gjessing, 2001); for instance a parameter region exists in which the first-passage-time distribution follows a non-monotonic behaviour.

The leaky competing accumulator (LCA) model (Usher & Cohen, 1999; Usher & McClelland, 2001) can be described as a multi-dimensional generalization of the OU-process. The LCA model was developed to model interactions between a number of alternative categories (or choices). Though the model was not directly targeted at modelling identification accuracy as a function of exposure duration we find no reason why this should not be attempted. On the other hand the LCA model has nevertheless been aimed at modelling a phenomenon such as reaction time, which besides time for perception also includes time for motor action (Usher & McClelland, 2001). The LCA model assumes an activation node A_x for each visual category, x .

$$dA_x = \left[\mu_x - \lambda_L \cdot A_x + \alpha \cdot F_x - \beta \sum_{z \neq x} F_z \right] \cdot dt + \sigma_x \cdot dw_x$$

The node A_x is activated by the external input, μ_x , when activated the node re-excites itself with $\alpha \cdot F_x$, where F_x is an activation function (e.g. a sigmoid function) that depends on A_i . Between nodes there is lateral inhibition, for instance the lateral inhibition from node A_z to any other node is $\beta \cdot F_z$. This means that when a node is activated it will inhibit all other nodes. Together with an exponential leakage term, λ_L the lateral inhibition helps balance the activation levels. We note that α , β and λ_L are all global parameters, meaning that in the model they do not depend on the specific node. An early version of the LCA type of model was provided in (Usher & Cohen, 1999). Later the model was slightly altered and substantiated in (Usher & McClelland, 2001, Bogacz et al., 2006)

A linear version of LCA also exist; this version which is more tractable mathematically, appears when the activation function F_x equals the activation A_x for all categories, x . When only one category is considered this linear version of LCA reduces to the Ornstein-Uhlenbeck process, as α and λ_L are both linear in the activation A_x

When considering identification in whole and partial report, then TVA actually describes the performance in these multi-element displays as a function of performance in single-element displays. A difference between single- and multi-element displays is that in the later case performance is also limited by VSTM storage capacity. Though the storage capacity is usually estimated to lie between 3 to 4 elements this ‘magical number’ has been vividly debated (Alvarez & Cavanagh, 2004; Cowan, 2000; Luck & Vogel, 1997; Miller, 1956). In TVA storage capacity is assumed to be limited, but nonetheless varying from trial to trial. As an example (Bundesen, 1990) assumed a mixture model for VSTM so that in some trials the capacity would be 4 and in the remaining trials the capacity would be only 3. Though this type of model fitted most of the data, it was unable to explain situations when 5 targets were reported, and so the model was not perfect. The mechanism that determines the storage capacity is not well understood, however a limited storage capacity does nevertheless seem as a sign of some inhibitory interaction between objects. In relation to the stochastic accumulator

models the concept of competition has been suggested (Deco & Rolls, 2004; Deco & Zihl, 2004; Usher & Cohen, 1999; Usher & McClelland, 2001) to account for such type of inhibitory mechanism. For this reason, we shall later study (in Section 6.3) whether a stochastic accumulator model such as the LCA model that we presented above might offer a possible explanation of the limited but varying storage capacity.

PART II:
METHODS AND RESULTS

5 Research question

Applying TVA to analyse accuracy in whole and partial report can be very informative for diagnostic purposes, for instance the procedure can be used to assess attentional deficits (such as visual neglect or simultanagnosia) following brain lesions (Duncan et al., 2003, 1999; Habekost & Bundesen, 2003). Further as we have already mentioned whole and partial report experiments can provide correlates of the patient state under a number of cognitive disorders (Habekost & Starrfelt, 2008) including Alzheimer's disease and Huntington's disease (Bublak et al., 2006; Bublak et al., 2009).

From whole and partial report experiments it is possible to infer knowledge about subject specific parameters such as the minimum exposure duration (the offset), μ that can lead to identification, the processing speed, C and the storage capacity, K . These cognitive parameters are often affected if a disease or abnormal condition is affecting the visual system, and therefore for diagnostic purposes it is important that we are able to estimate them precisely. Clearly whole and partial report experiments themselves are not enough to allow us to estimate cognitive parameters; we also need a model such as the theory of visual attention. TVA describes accuracy in whole and partial report experiments as a function of accuracy in single-object report, however few studies exist that have actually investigated accuracy in single-object report; for an exception see (Bundesen & Harms, 1999). Therefore the *principal aim of our research* is to provide a more detailed account of accuracy in single-object report. Single-object report is modelled by the so-called psychometric function, which relates accuracy to exposure duration. Traditionally TVA has assumed an exponential psychometric function; however we found no reason why other psychometric functions would be excluded as candidates. A more suited candidate for the psychometric function would improve model accuracy in single-object report. Further, as it is mathematically possible to integrate any psychometric function into the theory of visual attention (as we have already explained in Section 2.2), such psychometric function could potentially also improve the ability of TVA to account for accuracy in whole and partial report experiments. Also, providing a massive datasets with many trials our data might also allow us to analyse other aspects of data, such as how confusion develops as a function of exposure duration.

The thesis treats seven related research themes (In Sections 6.1-6.7):

- 1) *Single-letter report*: is it reasonable from a psychometric point of view to expect a minimum effective exposure duration (an offset), μ that is constant? Is it reasonable from a neural point of view to expect a processing capacity, C that is constant over time? In summary could a more optimal psychometric function than the exponential be identified? (In Section 6.1)
- 2) *Whole and partial report*: is it possible to use assume other psychometric functions than the exponential, when modelling accuracy in whole and partial report experiments? (In Section 6.2)

- 3) *Storage capacity*: related to the fact that (Shibuya & Bundesen, 1988) implements a storage capacity, K that varies between trials (by using FIRM as a mixture model), how could a trial-based variation in storage capacity be neuro-computationally plausible? (In Section 6.3)
- 4) *Stimulus contrast*: related to the fact that TVA does not advice that a particular stimulus contrast level should be used with whole and partial report, what would happen if stimulus contrast is altered, does the form of the psychometric function then change? (In Section 6.4)
- 5) *Stimulus identity*: should we expect that the form of the psychometric function is affected when performance is averaged across the individual stimulus identities? (In Section 6.5)
- 6) *Confusion*: does the pattern of confusion between stimulus identities change systematically as a function of exposure duration? (In Section 6.6)
- 7) *Guessing rates*: are guessing rates affected by previously reported or previously presented stimuli and do guessing rates vary over time? (In Section 6.7)

6 The research themes

In Sections 6.1-6.7 we treat each of the seven research themes presented in Chapter 5. Further Section 6.8 contains a summary of our results related to our principal research question namely whether it is possible to indicate a closer to optimal psychometric function (than the exponential) to use in connection with the modelling of identification accuracy in single-object report and possibly also in connection with modelling (using TVA) of identification accuracy in whole and partial report.

6.1 *Single-letter report: the psychometric function*

The exponential psychometric function used in TVA (Bundesen, 1990) assumed an offset, μ before visual identification could take place, and also the function assumed that processing capacity, C (which in single-object report equals the hazard of identifying the target) was constant over time. Other psychometric functions might not adhere to these rather strict assumptions, and so here we investigate if other psychometric functions are better suited for modelling data from single-letter report experiments.

We conducted a psychophysical experiment in which we investigated visual letter identification as a function of exposure duration (A. Petersen & Andersen, 2009). In our experiment (Experiment 1) three subjects each completed 54,080 trials in a 26-Alternative Forced Choice procedure. On each trial, a single randomly chosen letter (A-Z) was presented at the centre of the screen. Exposure duration was varied from 5 to 210 milliseconds.

The letter presented was followed by a so-called pattern mask. These pattern masks were randomly generated from trial to trial. The idea behind the mask that we used was to obtain a mask that would ideally affect the 26 stimuli from A to Z with an equally strong masking effect. Masks were rectangular and had fixed dimensions so that they would just cover any of the stimuli images. It has been shown that character identification is dependent on frequency channels (Oruç, Landy, & Pelli, 2006), and so we aimed at having masks with a frequency spectrum that was close to the average frequency spectrum of the stimuli images. If the subject is able to learn the appearance of the mask, there is a potential risk that he might also learn how to ignore it, and thereby increase his identification accuracy on the long run (Wolford, Marchak, & Hughes, 1988; Beeck, Wagemans, & Vogels, 2007). In order to avoid this potential effect of learning the mask we decided to use masks that were phase randomized from trial to trial, so that the appearance of the mask was never the same from trial to trial. Further, the mask images were subjected to a threshold so that the mask images would contain just as many foreground pixels as the average of the stimuli images. Compared to any stimuli image, a random binary mask image covered approximately the same area of the screen, had approximately as many foreground pixels and had approximately the same frequency spectrum. The phase content of the mask images was however

randomized from trial to trial. A detailed description of the procedure used to generate the masks is available in Appendix A in (A. Petersen & Andersen, 2009).

In (A. Petersen & Andersen, 2009) we describe six psychometric functions, which, for various reasons, are plausible candidates for describing letter identification as a function of exposure duration. In Bundesen's original TVA he used the exponential distribution with an offset, μ as the psychometric function. We compared the exponential, the gamma and the Weibull psychometric functions, all of these having a temporal offset included, as well as the ex-Gaussian, the log-logistic and finally the squared-logistic, where the later is a psychometric function we believe have not been described before. The 3-parameter Weibull, gamma and ex-Gaussian distributions all contain the 2-parameter exponential distribution as a special case, the 2-parameter log-logistic and the 3-parameter squared-logistic distributions do not.

All fits were conducted based on maximum-likelihood based procedures. The log-logistic and the squared-logistic psychometric functions fit well to our experimental data, this can be seen in Figure 1, which shows residuals plotted as a function of exposure duration.

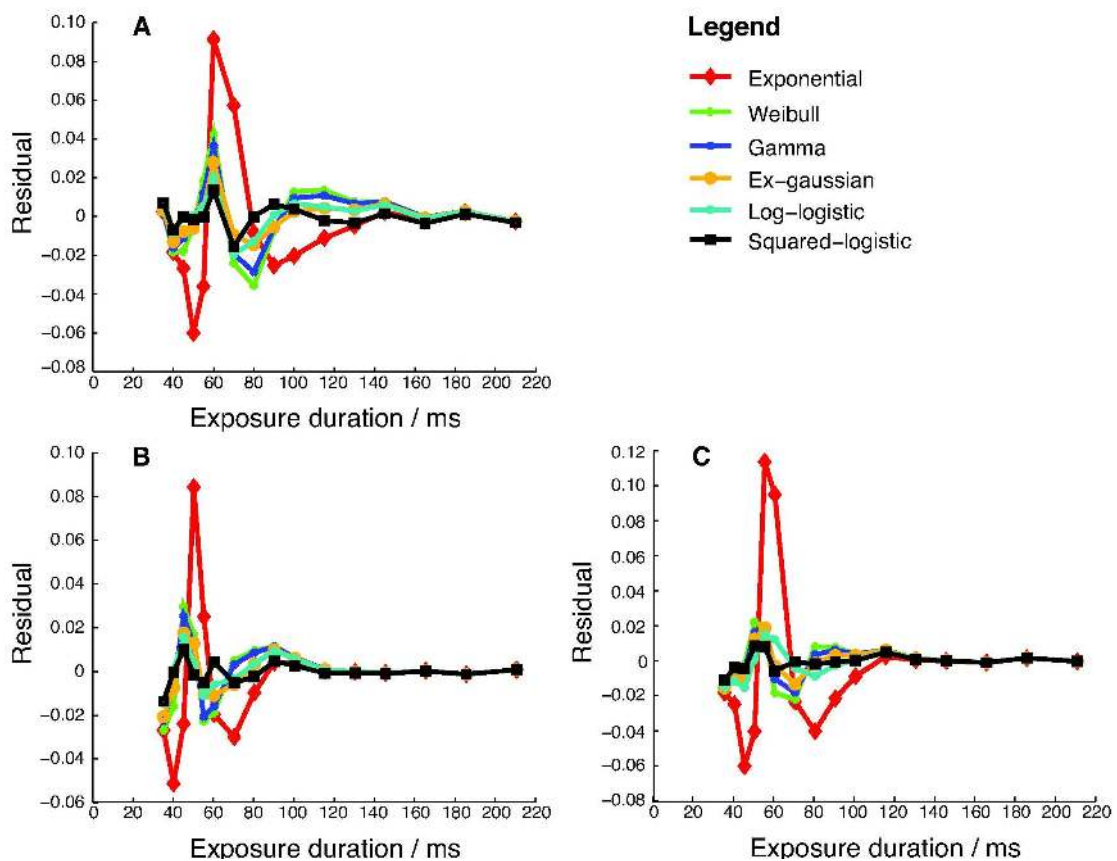


Figure 1: Residuals plotted as a function of exposure duration from Experiment 1. Errorbars – too small to be distinguished clearly – show the standard error of the mean. There is one graph for each subject: a) AP, b) MH and c) MK.

For all three subjects the squared-logistic provided much better fits than the exponential. This is evident both in Figure 1 and in Figure 2, which shows proportion correct together with the fits of the exponential and the squared-logistic psychometric functions.

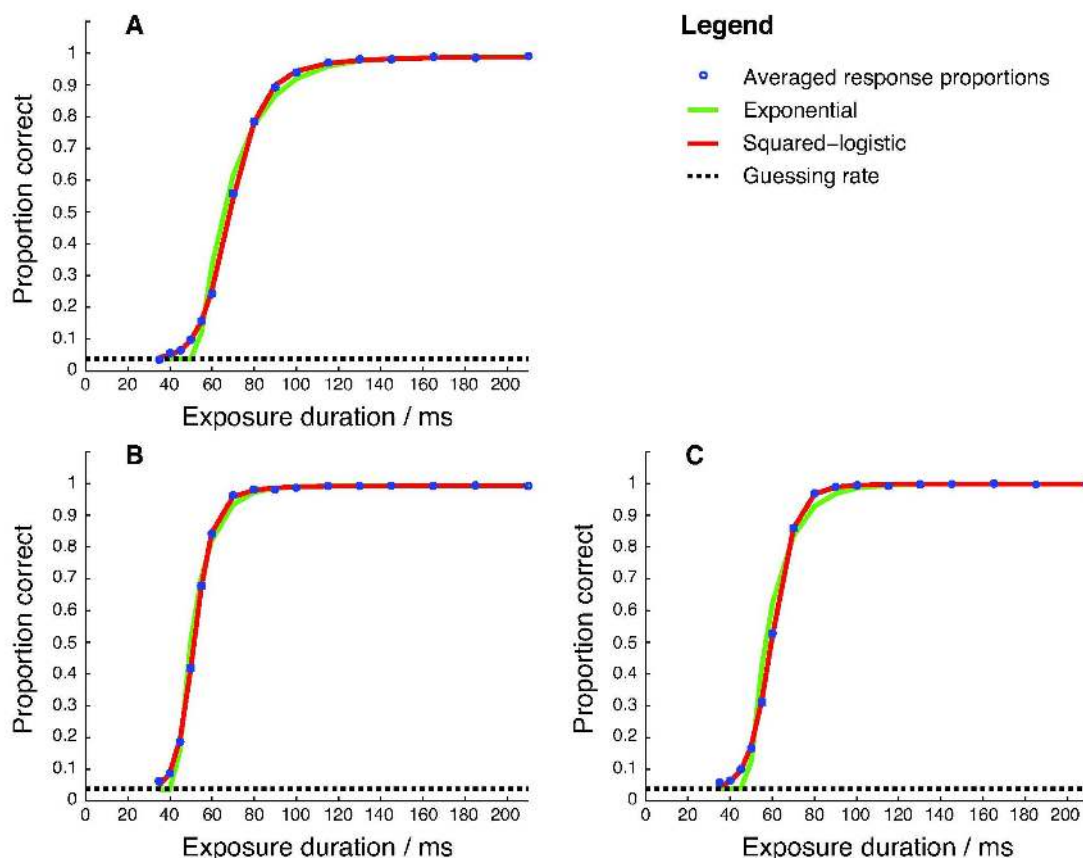


Figure 2: Proportion correct averaged over letter identities. Error bars show the standard error of the mean. The fit of the exponential and the squared-logistic psychometric functions are shown as well. There is one graph for each subject: a) AP, b) MH and c) MK.

Rather than plotting residuals such as in Figure 1, an alternative way to inspect the fits of the six psychometric functions is to plot their hazard functions (Figure 3). The hazard function characterizes the conditional probability density of encoding the object given that this has not previously happened. In Figure 3 we see the empirical hazard function, which we calculated from the data (Van Zandt, 2002), plotted along with the hazard function of the six psychometric functions, when these were fit to the empirical data. It is seen that the empirical hazard rate seems to evolve in a non-monotonic fashion as a function of time. The exponential distribution does not have a non-monotonic hazard function, and nor do the Weibull, the gamma or the ex-Gaussian distributions; the log-logistic and the squared-logistic distributions, on the other hand, both show a hazard function that is non-monotonic.

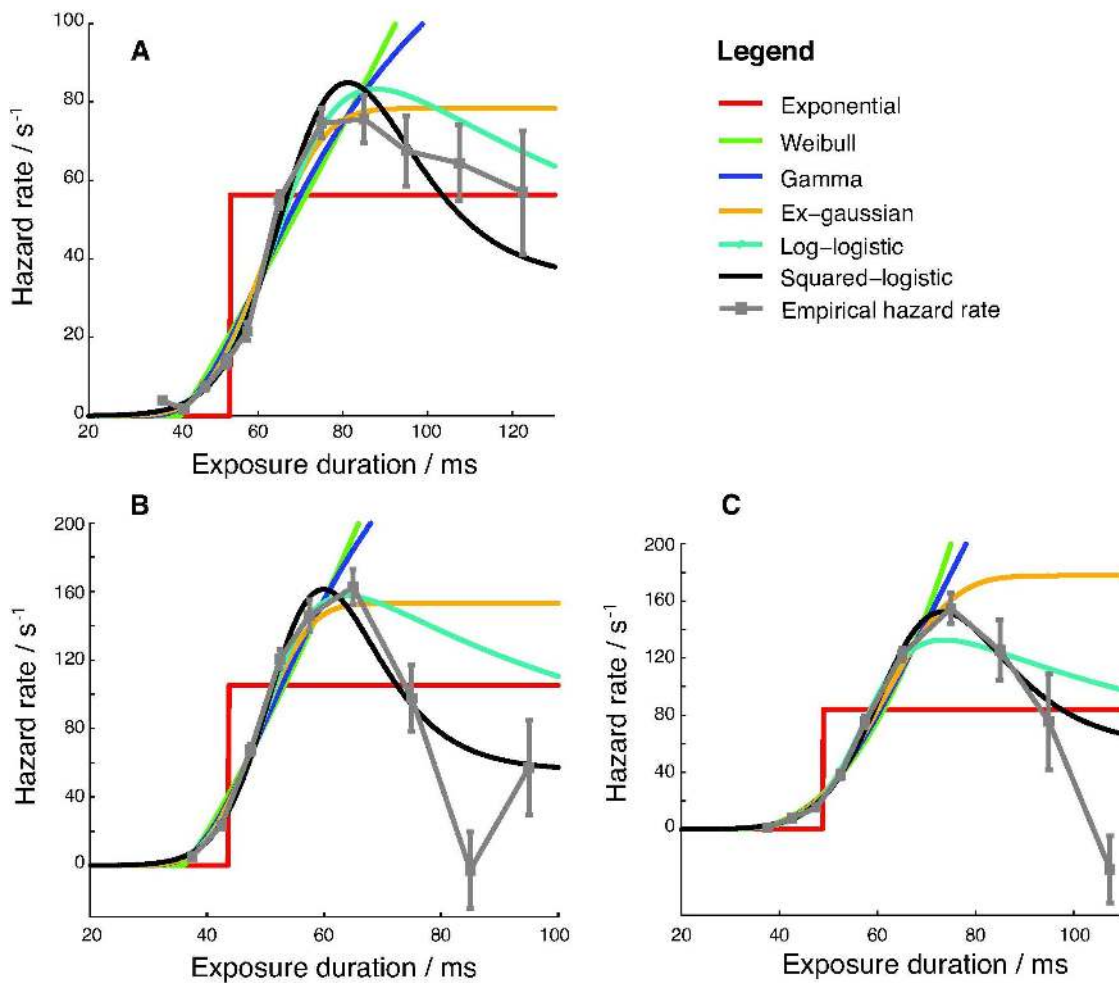


Figure 3: The hazard function plotted against exposure duration. Grey squares represent the empirical hazard function, which was estimated directly from the data. Error bars indicate the standard deviation as estimated by a bootstrap procedure. The hazard functions of the various psychometric functions are displayed as coloured lines. There is one graph for each subject: a) AP, b) MH and c) MK.

6.2 Whole and partial report: the generalized FIRM equations

Originally TVA was formulated based on the assumption of an exponential psychometric function (Bundesen, 1990). To allow for other psychometric functions than this, the fixed-capacity independent race model (FIRM) equations that implement the storage capacity limit in TVA, have to be generalized. As we have already mentioned in Section 2.2 a contribution from our work is that we generalized the FIRM equations. The generalized FIRM equations which can be found in Appendix B in (A. Petersen & Andersen, 2009) allow the insertion of any desired psychometric function into TVA, and therefore TVA does no longer have to assume an exponential psychometric function. Therefore, with the generalized FIRM equations it has become possible to investigate if other psychometric functions generate better fits to data from whole and partial experiments than does the exponential psychometric function.

With the generalized FIRM equations, we inserted each of the six psychometric functions into TVA. The six models were fitted to the data for each of the two subjects in (Shibuya & Bundesen, 1988). Figure 4 shows the cumulative score distributions for these two subjects. The score is the number of correctly reported targets in a given trial. The score distribution is the distribution of correctly reported targets seen over trials having the same experimental conditions. The experimental conditions are defined by the number of targets, T and the number of distractors, D in the display as well as the exposure duration. The cumulative score distribution shows the probability of reporting at least j targets correctly. A careful visual inspection of Figure 4 reveals that the squared-logistic based model fits closer than the model based on the exponential psychometric function; this is true for both subject MP (Figure 4.a) and subject HV (Figure 4.b).

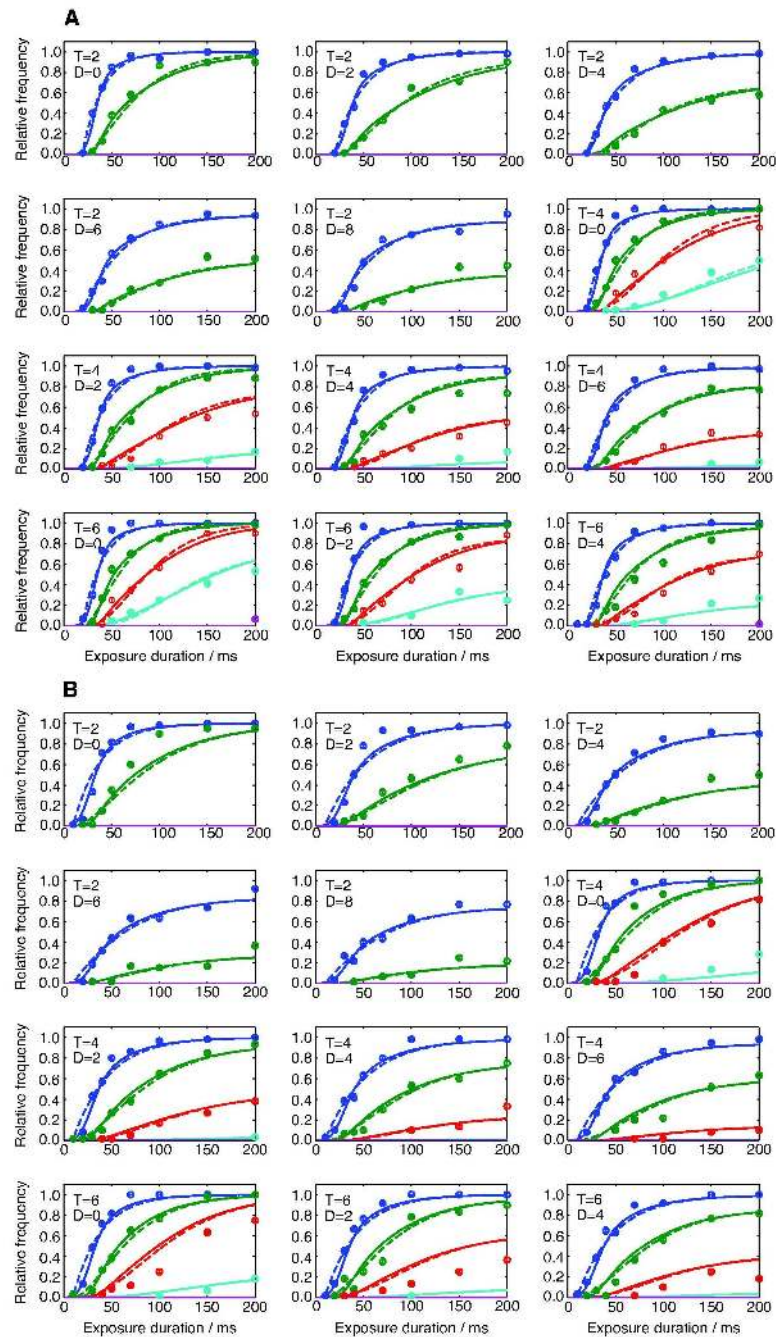


Figure 4: Cumulative score distributions from a whole and partial report experiment (Shibuya & Bundesen, 1988). Circular markers correspond to the proportion of scores of j or more (correctly reported targets) showed as a function of exposure duration. The legend for the scores is blue: $j=1$, green: $j=2$, red: $j=3$, turquoise: $j=4$ and violet $j=5$. Each graph shows data for a certain combination of the number of targets, T , and the number of distractors, D . The dotted lines represent the fit of the exponential psychometric function inserted into TVA. Solid lines represent the fit of the squared-logistic psychometric function inserted into TVA. There are two sub-figures, one for each subject: a) MP and b) HV.

6.3 The varying storage capacity of the visual short-term memory

At a first glance it seems strange why VSTM storage capacity, K would vary from trial to trial as assumed in TVA. What neural mechanisms could cause such behaviour, and is it plausible that capacity is determined as a side product of some competitive interaction between the neural representations of the various objects in the visual field? A model for multi-object representation and interaction is the LCA model (Usher & Cohen, 1999; Usher & McClelland, 2001), which although it is still quite simple, was constructed to mimic interactions between ensembles of neurons. In a first attempt (A. Petersen et al., 2009) to integrate TVA (Bundesen, 1990) with the LCA model (Usher & Cohen, 1999) the model that we suggested was a slightly modified version of the LCA model. The model was defined as:

$$\frac{dA_x}{dt} = -A_x + \alpha F(A_x) - \beta \sum_{z \neq x} F(A_z) + \gamma G(v_x)$$

Modifications compared to the LCA model described in Chapter 4 were: 1) A unity leakage term was assumed 2) the Gaussian noise term was removed from the equation and 3) the input term was replaced with a Poisson spike train, G . The amplitude of the input is determined by the scaling factor, γ . The rate of arrival of the Poisson input spikes supporting the categorization of element x was defined as the hazard rate, v_x of encoding element, x . Hazard rates (processing rates) were determined according to traditional TVA principles (see Section 2.1). More details can be found in (A. Petersen et al., 2009).

In order to investigate if an LCA type of model can be used for determining the storage capacity limit in TVA, we evaluate the new model's ability to fit experimental data from a classical whole and partial report study (Shibuya & Bundesen, 1988). The results from the fit can be seen in Figure 5 which shows again the cumulative score distribution for subject MP (same empirical data as shown in Figure 4.a above). As we mentioned previously the cumulative score distribution describes the probability of reporting at least j targets correctly.

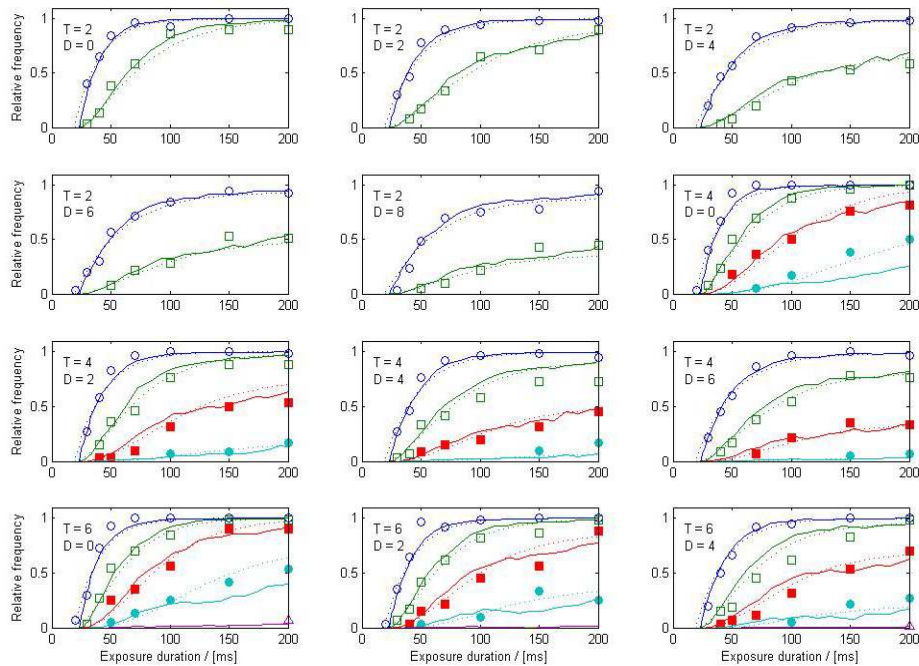


Figure 5: Cumulative score distributions for subject MP in (Shibuya & Bundesen, 1988). Probability of correctly reporting at least $j=1$ target (blue, open circles), $j=2$ targets (green, open squares), $j=3$ targets (red, closed squares), $j=4$ targets (cyan, closed circles) and $j=5$ targets (magenta, open triangles). Empirically found values are plotted with symbols as markers. The dotted lines represent the fit by Shibuya & Bundesen (1988). Solid lines represent the performance of our neural network model. T and D denote the number of targets and distractors presented, respectively.

Figure 5 gives a qualitative indication that using LCA to determine storage capacity yielded quite similar performance, as when one used the fixed-capacity race model, which has traditionally been a part of the TVA framework (Bundesen, 1990). A difference between Bundesen's TVA and our new model was that the later was able to predict extreme scores of 5, indicating that storage capacity varies more broadly with the new model.

6.4 Individual identities: is averaging okay?

In the previous section we looked at performance when this was averaged across the different letter identities. Averaging was performed despite the fact that there is no obvious *a priori* reason to assume that the psychometric function preserves its shape when averaged across several stimuli in an identification task. In fact, the psychometric function for identification of individual letters as a function of contrast was investigated in (Alexander, Xie, & Derlacki, 1997). This study demonstrated a significant parametric dependency of the psychometric function on letter identity when investigating identification for 10 different Sloan letters. Though we vary exposure duration rather than contrast, we think it is reasonable to question whether averaging over letter

identities may affect the shape of the psychometric function. Therefore, we also fit the 6 psychometric functions to the data *without* averaging across letter identity.

In Figure 6 we illustrate the parameter differences between letters by showing parameter histograms as well as parameter scatter plots; data for all three subjects are shown. Under the assumption that the psychometric function obtained by averaging over letter identities is the true psychometric function we applied bootstrapping (Efron & Tibshirani, 1994) to check if the variance in the parameters of the psychometric functions for the individual letter identities can be ascribed entirely to an effect of random sampling or if it needs also to be ascribed to some systematic effect of letter identity. The bootstrapping consisted of fitting the psychometric function to 200 random re-samples of the data (averaged over letter identities) and then calculating the standard deviation for each of the model parameters based on the 200 model fits. The ovals in Figure 6 demark the bootstrap estimated confidence region to which we would expect 95 % of the letters, placed according to their individual parameters, to be located. Clearly many letters are located outside the ovals. This shows that the model parameters vary significantly with letter identity.

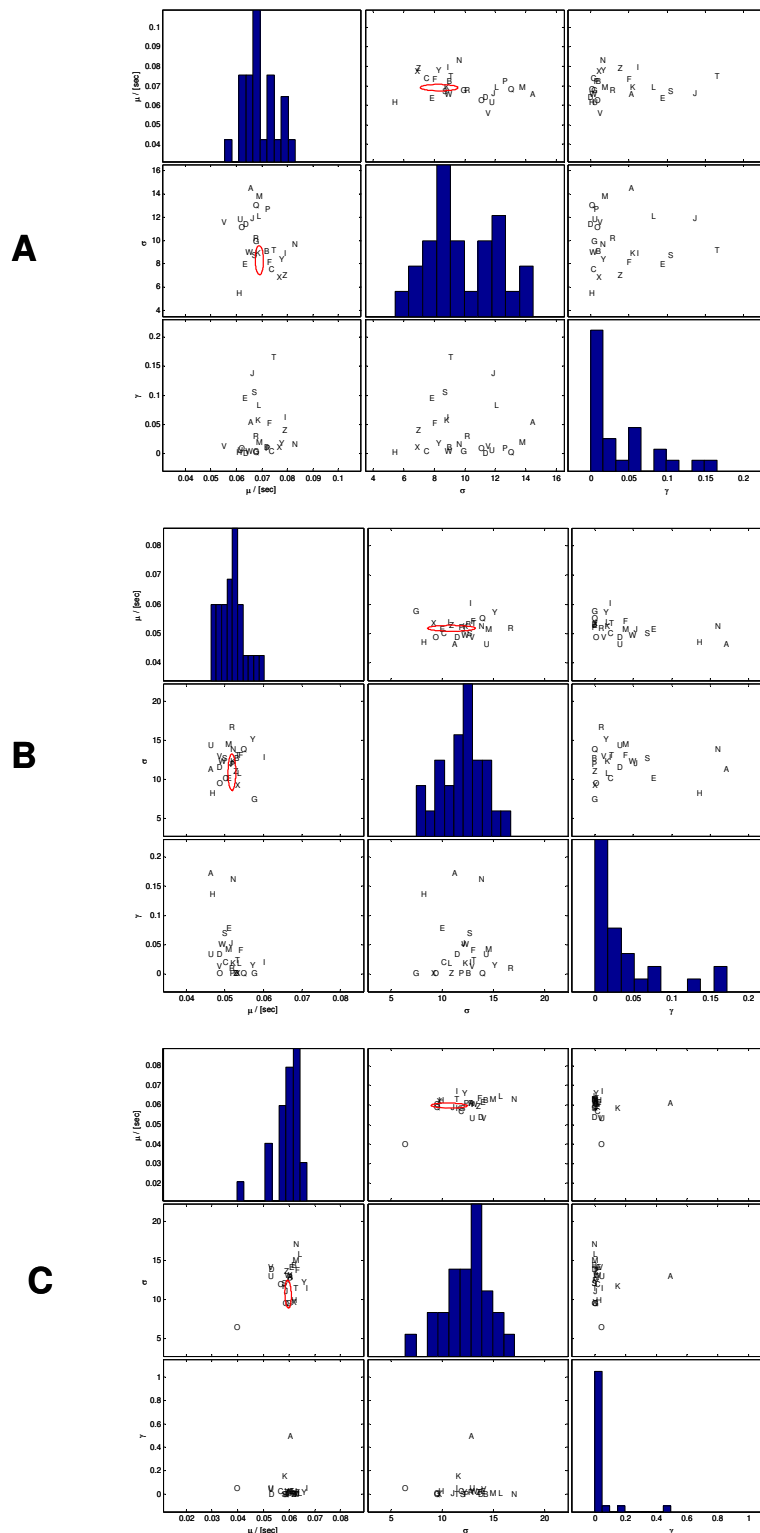


Figure 6: Scatter plots and histograms for the parameters of the log-logistic model fitted to the data from Experiment 1. In the off-diagonal windows we see parameter pairs plotted against each other, while in the diagonal windows we see the parameter histograms. Each letter from A to Z was fitted individually giving each letter its own set of model parameters. There is a graph for each subject: a) AP, b) MH and c) MK.

6.5 Confusion: does it develop with exposure duration?

Our dataset provides a unique opportunity to investigate confusion between letter identities. In Figure 7, we show the logarithm of the posterior mean for reporting the letter O for the different stimulus letters and the different exposure durations. This Bayesian approach (Ghosh, Delampady, & Samanta, 2006) is especially appropriate because for a large number of trials the posterior mean most often approach the maximum likelihood estimate; however the posterior mean has the advantage that it is always positive, this provides us with an opportunity to go into log-space. Seen across subjects, when the stimulus is actually an O the posterior mean increases with the stimulus exposure duration. For the stimulus letters C, G and Q the posterior mean first increases, then decreases as a function of exposure duration. This later pattern is especially evident for the letter Q. For subject MK it is particularly striking how all roundish letters: B, C, D, G, J, P, Q, R, S and U are the ones that most often are perceived to be an O. This shows that some letters are more confusable than others and that confusability is likely to be affected by the features shared between letter identities.

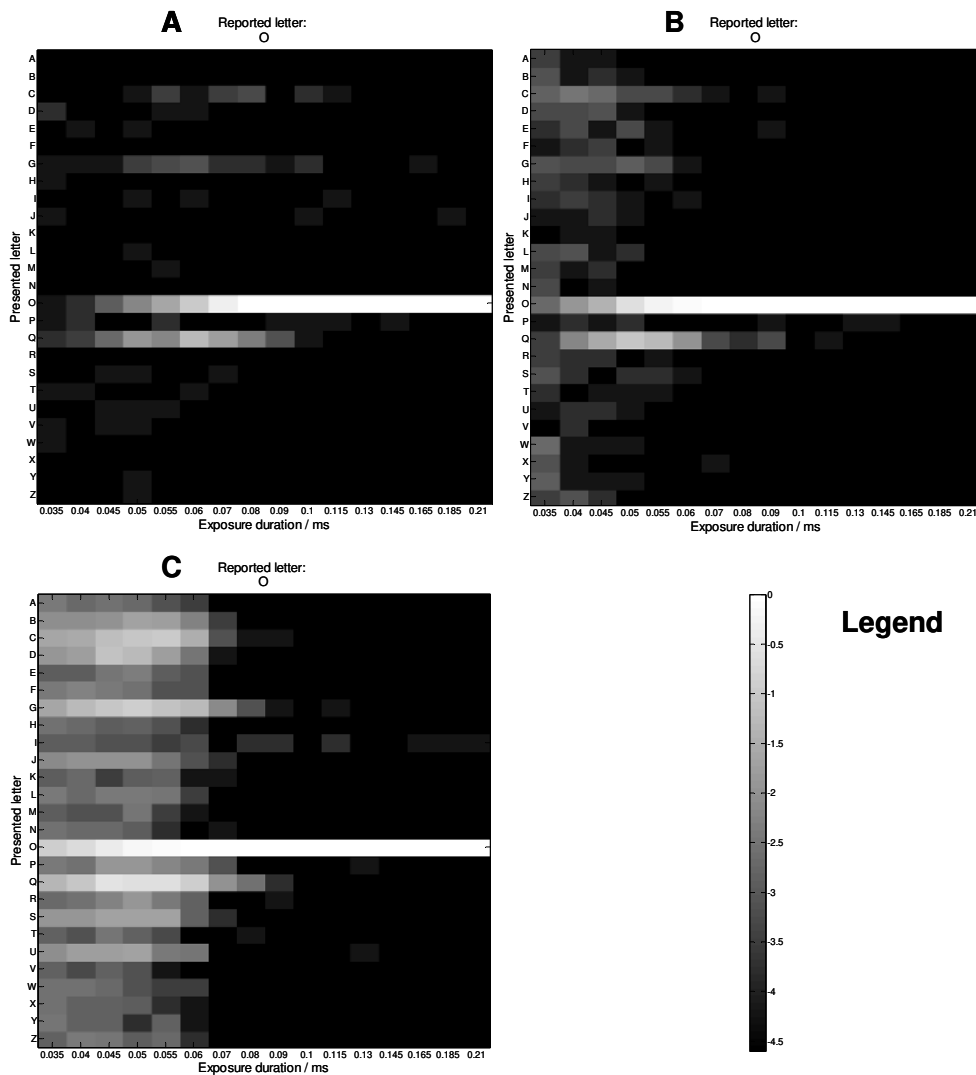


Figure 7: Log posterior mean of reporting an O shown as function of exposure duration and identity of the letter presented. The log-posterior mean was calculated assuming a uniform prior distribution. There is a graph for each subject: a) AP, b) MH and c) MK.

6.6 Contrast: does it affect the form of the psychometric function?

In (A. Petersen & Andersen, 2009) we conducted a follow-up contrast experiment (Experiment 2) to test the ability of six psychometric functions (already mentioned in Section 6.1) to fit single-letter identification data at different stimulus contrast levels. Experiment 2 was quite similar to Experiment 1 (described in Section 6.1). In Experiment 2 one subject (AP) completed 36,400 trials in a 26-Alternative Forced Choice procedure that encompassed exposures at 11 different contrast levels. In Figure 8 we see how exposure duration and contrast level influence the proportion correct observed in Experiment 2. It is also seen how the squared-logistic psychometric function (which was the psychometric function that best characterised the data in Experiment 1) fits quite well to the data regardless of the contrast level used. This implies that the form of psychometric function is not dependent on the contrast level.

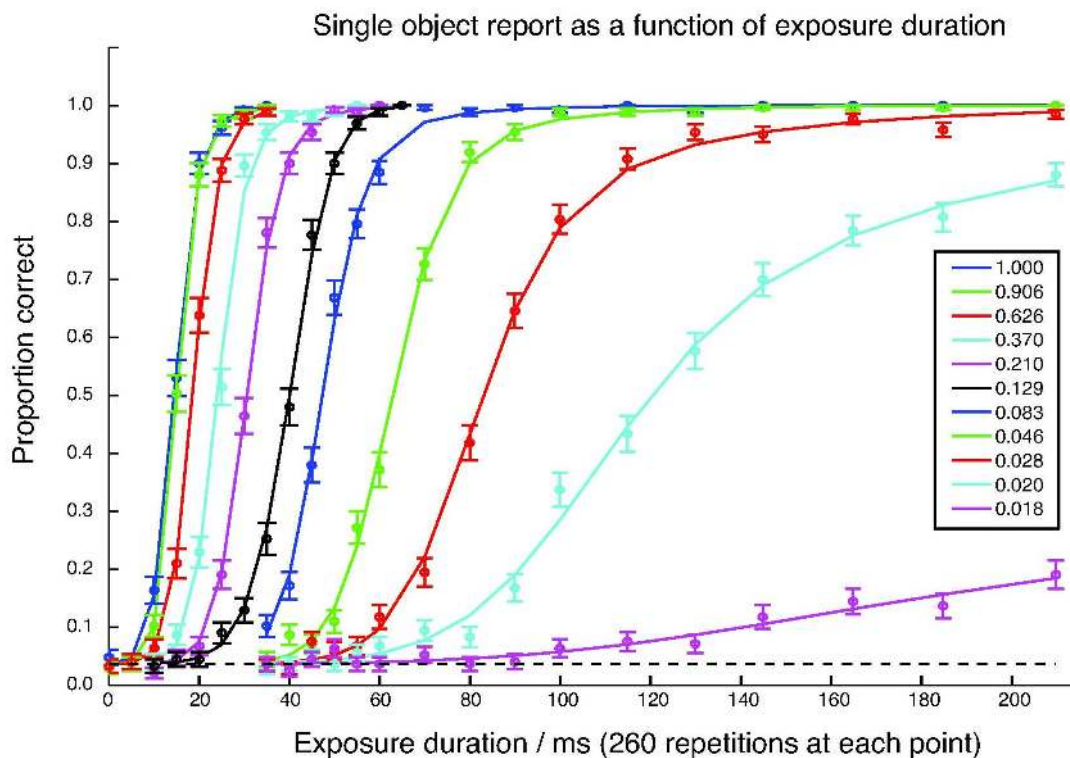


Figure 8: Proportion correct averaged over letter identities for subject AP in our contrast experiment. Error bars show standard deviation of the mean. Also shown (as lines) are the fits of the squared-logistic psychometric function. The legend shows the negative Weber contrast that was used.

In our experiments we varied both exposure duration and contrast, however our psychometric function is only defined as a function of the earlier. It is however unquestionable that the parameters of our psychometric function varies also with contrast. In Figure 9 we show how the estimated model parameters vary as a function of contrast. There seems to be a monotonic mapping between the three parameters of the squared-logistic psychometric function when the contrast is varied. With respect to the points corresponding to the fit to the negative Weber level of 0.129 these points seem to

deviate a bit from the general tendencies seen in the rest of the data points. The reason for this is unknown. This could perhaps be caused by a possible trading in parameters, for instance μ appear somewhat high, but clearly also σ and V appear quite high. Generally we see that when contrast is increased μ and σ go down whereas V goes up. For negative Weber contrast levels between 1 and 0.028 the relationship between μ and σ appears somewhat linear. In summary, there seems to be a systematic development in the model parameters as a function of stimulus contrast, still we have not derived expressions that describe the dependency of the psychometric function upon the contrast level.

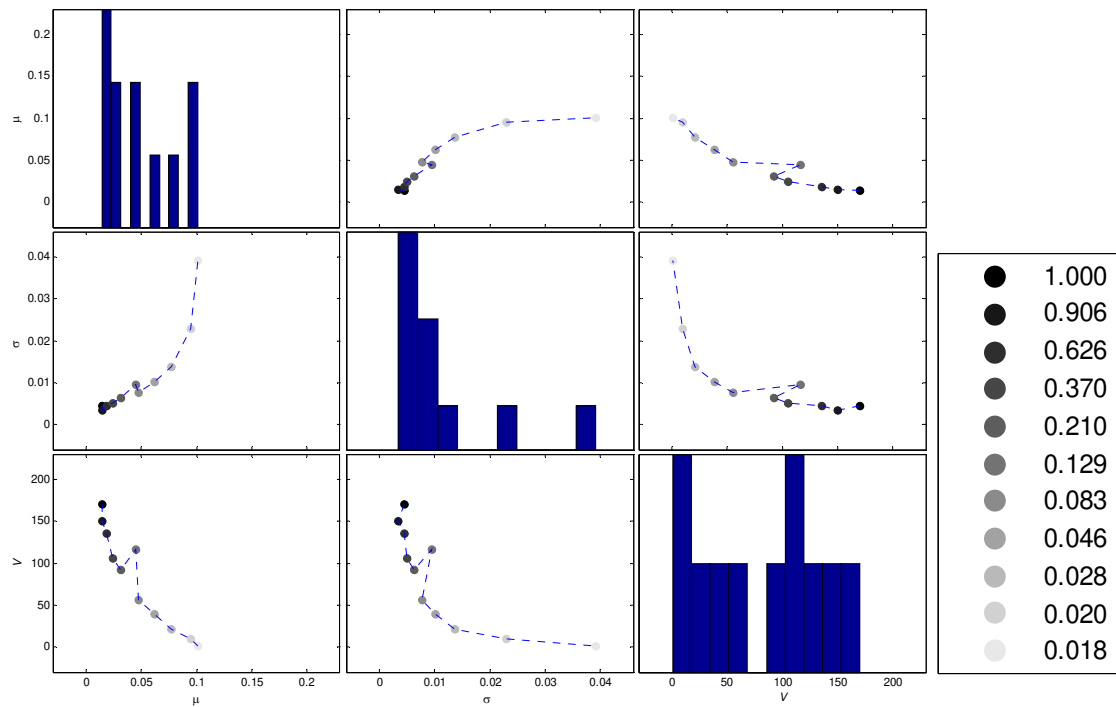


Figure 9: Development of parameters as a function of contrast for subject AP in the contrast experiment. The figure contains 9 windows. There are 11 point markers in each graph in the off-diagonal windows, one for each stimulus contrast level in the experiment; the darker the point marker, the larger the negative Weber contrast used, this is indicated in the legend. The diagonal windows show parameter histograms for the various contrast levels.

6.7 Guessing rates: what do they reveal?

When stimuli are presented in a sequential manner one might speculate whether the response in a given trial is influenced by earlier responses as well as stimuli earlier presented (Maloney, Dal Martello, Sahm, & Spillmann, 2005; Mozer, Kinoshita, & Shettel, 2007). The lag is defined as the number of trials between two sequential trials, per definition the first of two such sequential stimuli is said to be presented at lag-0. Let us define that the prior probability of a certain letter is the probability of reporting that letter not knowing *a priori* what letter is actually shown, the prior probability can easily be calculated from the data as the proportion of trials a certain letter is reported, averaging over all presentation conditions. From our dataset it is possible to investigate

how the prior probability of reporting a letter identity at a certain lag depends on the *reported* letter at 0-lag. Similarly we can investigate if the *presented* letter at 0-lag influences the reported letter at a later lag.

For the dependency of the reported letter at lag $n > 0$ on the *reported* letter at lag-0 Let us define that the prior probability of reporting the letter x at a certain lag n is denoted as: $P(X=x)$ and that the prior probability of *reporting* the letter y at lag-0 is denoted as: $P(Y=y)$.

Then the joint probability of reporting x at lag n and y at lag-0 is: $P(X=x \wedge Y=y)$
If we assume independence we must expect that

$$M(n)_{26 \times 26} := P(X=x \wedge Y=y)_{26 \times 26} / (P(X=x)_{26 \times 1} * P(Y=y)_{1 \times 26}) = 1_{26 \times 26}$$

Therefore if the elements in $M(n)$ are far away from 1 this would generally indicate that the letter reported at lag n is influenced by the letter reported at lag-0. The reason why elements would generally not be equal to 1 is because of variation due to random sampling. To test whether the outcome of Y influences the outcome of X we define a list of on-diagonal elements $L_{on}(n) = \text{diag}(M(n))$. The complementary list of off-diagonal elements we define as $L_{off}(n)$. Our aim is to investigate whether the letter reported at lag n is significantly influenced by the letter reported at lag-0. The statistical test that we use is a two-sample t-test (at 5% significance level) where we test whether the elements in $L_{on}(n)$ stem from a different distribution than the elements in $L_{off}(n)$.

Similarly, we also test for the dependency of the reported letter at lag $n > 0$ on the *presented* letter at lag-0. In this case we simply define $P(Y=y)$ to be instead the probability of the letter y being *presented* (rather than reported) at lag-0.

From Table 1 we can see that both the reported letter and the presented letter at lag-0 influence significantly the eight following letters that are reported.

Subject	AP	MH	MK	AP	MH	MK
Lag/Type	Report	Report	Report	Present	Present	Present
1	S	S	S	S	S	S
2	S	S	S	S	S	S
3	S	S	S	S	S	S
4	S	S	S	S	S	S
5	S	S	S	S	S	S
6	S	S	S	S	S	S
7	S	S	S	S	S	S
8	S	S	S	S	S	S
9	S	S	S	S	I	S
10	S	I	S	I	I	I
11	S	S	S	S	I	I
12	S	I	S	I	I	I
13	S	I	I	I	I	I
14	S	I	I	I	I	I
15	I	I	S	I	I	I
16	S	I	I	S	I	I
17	I	I	I	I	I	I
18	I	I	I	I	I	I
19	I	I	S	I	I	I
20	I	I	I	I	I	I
21	I	I	I			
22	I	I	I			
23	I	I	I			
24	I	I	I			
25	I	I	I			

Table 1: For each lag it is indicated whether the reported letter depends significantly (S) or insignificantly (I) on the letter reported at lag-0. Similarly, for each lag it is indicated whether the reported letter depends significantly (S) or insignificantly (I) on the letter presented at lag-0. It is possible to compare between subjects: AP, MH and MK and between dependency type: report and present.

We see that both reported and presented letter at lag-0 plays a significant role for all reported letters 8 lags ahead. The median time between presentations is *guesstimated* to be about 3 seconds. This means that a letter presented at lag 8 is presented approximately 24 seconds after the letter presented at lag 0. The sensory (iconic) memory is expected to fade away within a second (Sperling, 1960) whereas the VSTM memory trace is often assumed to last for around 30 seconds (Brown, 1958; L. R. Peterson & M. J. Peterson, 1959; A. D. Baddeley & Scott, 1971). For this reason we think our dataset offers evidence on how items, having been either partly or completely encoded into our short-term memory system, influence encoding of trailing items. Across subject it is seen that reporting a letter implies a longer-lasting influence on later reports than does simply being presented to a letter.

When estimating the psychometric function it is customary to assume that guessing rates do not vary over sessions. However, it might be the case that guessing rates do actually vary over sessions. Our dataset offers a unique chance to investigate this type of variation. It would be natural to assume that guessing rates must be close to the prior probability of reporting certain letter identities regardless of the identity of the letter actually presented. A difference is that where the guessing rates depend slightly on the

model assumed, the prior probabilities of the various letter identities do not. The reason for this is that the guessing rates are estimated together with the model chosen while the prior probabilities are calculated directly from the data. Therefore in Figure 10 we show the prior probabilities, and we can see how these vary quite systematically over sessions. This is perhaps most clearly seen for subject MK who over sessions shifts from reporting A's to reporting K's and O's. This shows that systematic variation in the prior probabilities (the guessing rates) over sessions can in some cases occur.

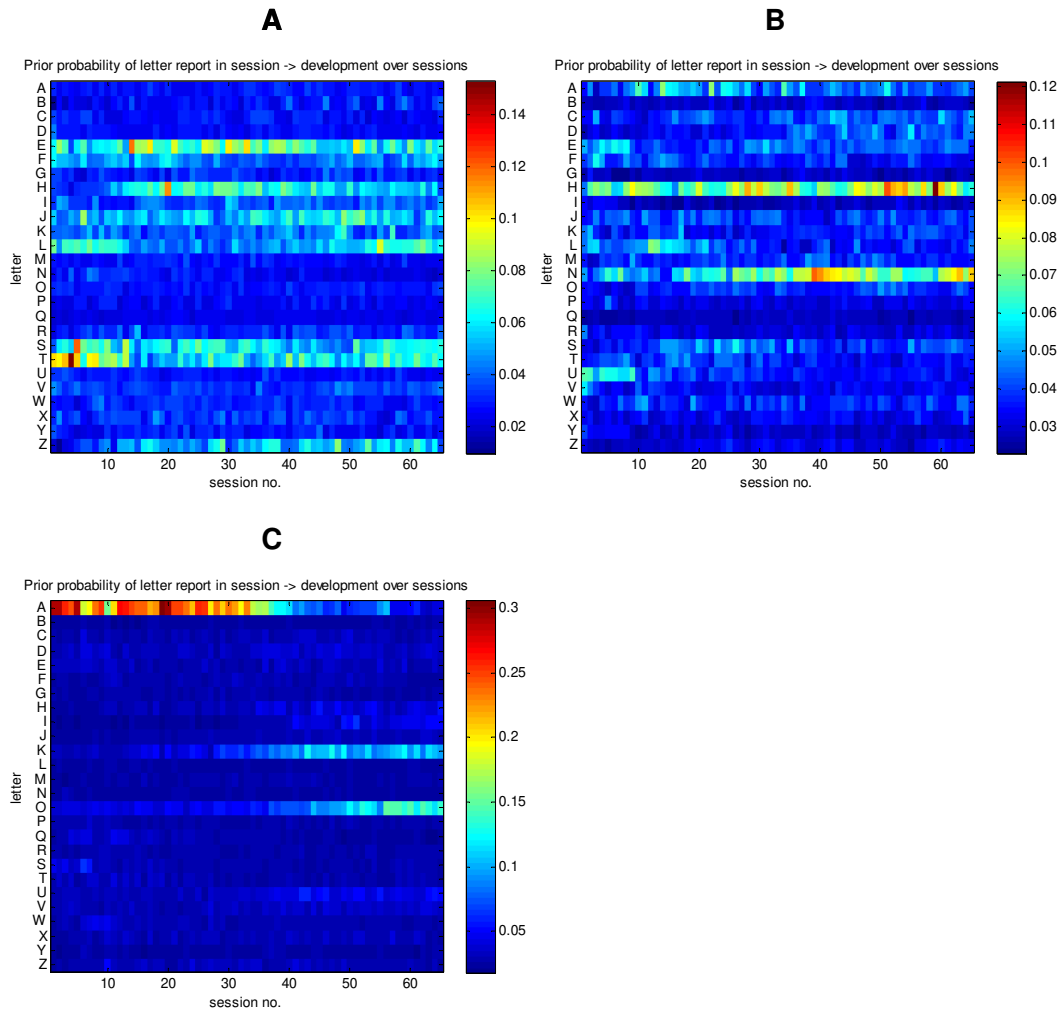


Figure 10: Prior probability for reporting each of 26 letter identities shown for each of the 65 experimental sessions. The prior probability is the probability of reporting a specific letter before it is decided what letter should be shown. There is a graph for each subject: a) AP, b) MH and c) MK.

6.8 Summary on our principal research question - the psychometric function

In the previous sections we explained how we used 6 different psychometric functions to fit data from original experiments as well as data from (Shibuya & Bundesen, 1988). We further fitted the same psychometric functions to data from a single-letter

experiment that included 3 subjects and involved 4000 trials per subject (Bundesen & Harms, 1999). In Figure 11 we show AIC (Akaike, 1974) and BIC values (Schwarz, 1978) for the six models that we tested as well as for the saturated model which is shown for reference purposes. AIC and BIC values are shown for the various datasets that we fitted, and for all of these we can see that the log-logistic and the squared-logistic psychometric functions are clearly better models than the exponential psychometric function.

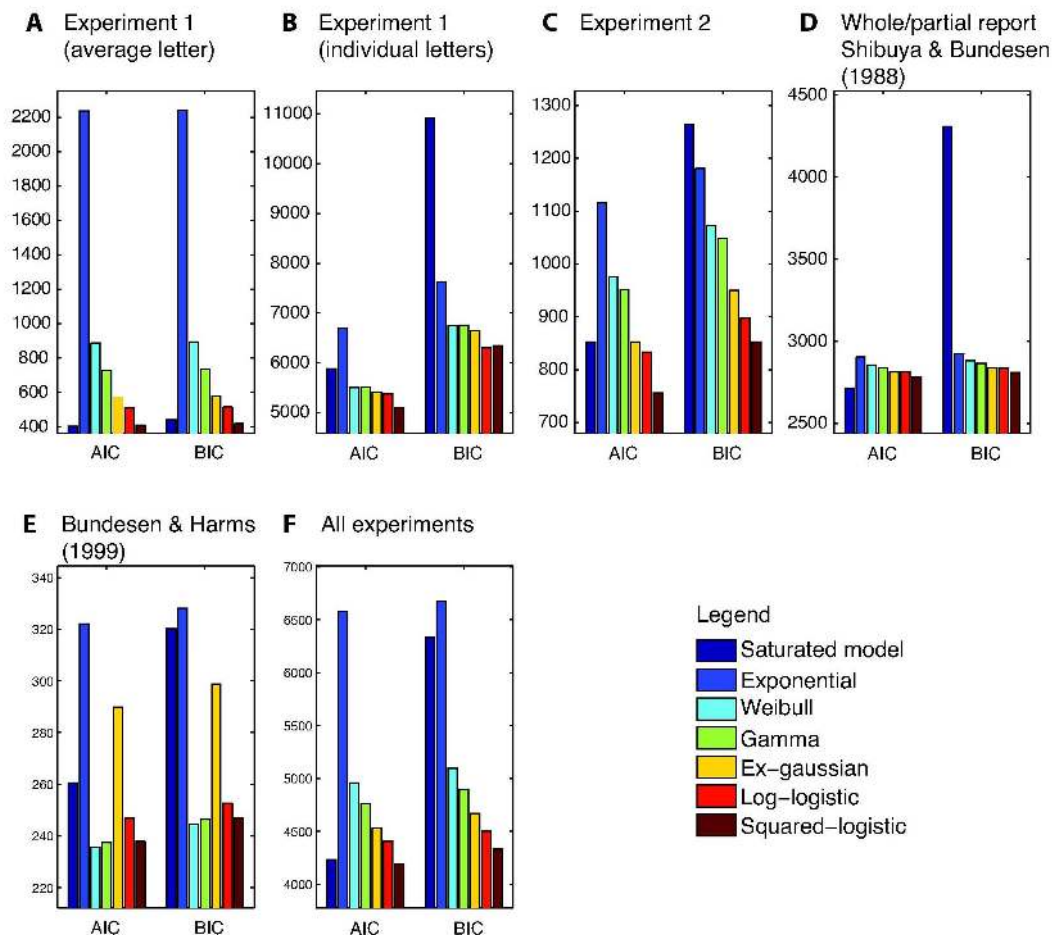


Figure 11: Model feasibility cumulated over all subjects. In the different graphs we see how successful the different psychometric functions are at modelling various datasets. The measures shown are AIC values (Akaike, 1974) and BIC values (Schwarz, 1978). In a) we see the results from fitting the function to the data from Experiment 1 averaged over letter identity. In b) we see the results from fitting the function to the data from Experiment 1 without averaging over letter identity. In c) We see the result from fitting to the data from the whole and partial report experiment in (Shibuya & Bundesen, 1988). In d) we see the results from fitting to the single-letter identification data from (Bundesen & Harms, 1999). In e) we see the result from fitting the functions to the data from Experiment 2. Finally in f) we see AIC and BIC measures cumulated over all datasets from the various experiments, including only the fit to the average single-letter data (i.e. not the fit to the individual letters) for Experiment 1.

**PART III:
CONCLUSION**

7 Conclusion

A review of TVA led to a series of research questions closely related to the psychometric function in TVA. In order to answer these questions we conducted extensive original psychophysical experiments (single-letter report), developed a new psychometric function (the squared-logistic) as well as we generalized an existing model of short-term memory capacity (the FIRM model) allowing us to integrate our new psychometric function into this model. The generalized FIRM model thereby allows us to address not only identification accuracy in single object report, but also identification accuracy in whole and partial report experiments.

Single-letter

In (A. Petersen & Andersen, 2009) we investigated visual letter identification as a function of exposure duration in order to study the shape of the psychometric function. Our experimental design reflected the fact that we deliberately aimed at having many trials per subject. In this way we could avoid averaging over subjects, which would possibly have affected the shape of the psychometric function. For practical reasons this only allowed us to include three subjects in the experiments. Fitting to original data from our comprehensive single-letter identification experiment (Experiment 1), the squared-logistic and the well-known log-logistic were shown to be the most optimal psychometric functions that we tested. Worth noticing both of these can be parameterized so that they have a non-monotonic hazard function. The squared-logistic, a psychometric function we found no previous accounts for, was developed with neurobiological motivation from NTVA (Bundesen & Habekost, 2008; Bundesen et al., 2005) and single-neuron studies by (Albrecht et al., 2002).

In a psychological experiment such as single-object report we can not directly observe how the stochastic accumulation of activation takes place in the respective areas of the brain, however as we have suggested in Section 3.1 the proportion correct could reveal information about the first-passage-time of the underlying stochastic process, and so from this perspective our single-letter identification experiment would allow us to narrow down the field of model candidates that might be considered. In this respect our investigation seems to suggest that the first-passage-time distribution of the underlying process for character identification as a function of exposure duration has a non-monotonic hazard function.

Whole and partial report

In addition to the original squared-logistic psychometric function, model synthesis also comprised generalizing the FIRM equations of TVA, which allowed us to assume any desired psychometric function when applying TVA to fit data from whole and partial report experiments. Inserting each of the psychometric functions into the Theory of Visual Attention (Bundesen, 1990), we fitted each of these models to data from whole and partial report type of experiments. We found that the closest to optimal

psychometric functions, also for modelling whole and partial report, where the squared-logistic and the log-logistic.

Capacity

We showed that a model mimicking interactions between neuron ensembles is capable of explaining the variation in the storage capacity limit of the visual short-term memory. This was demonstrated when we modelled a whole and partial report data set using TVA to provide the input to an LCA type of model (A. Petersen et al., 2009).

Contrast

A confound to the generality of our psychometric function would be if the type of psychometric function depended on stimulus contrast. Therefore, we investigated accuracy in single-letter identification as a function of exposure duration at 11 different contrast levels (Experiment 2). We then fitted each of the six psychometric functions to the dataset. Our results showed that even for this broad range of contrast levels the squared-logistic and the log-logistic were the most optimal and the exponential the least optimal of the six psychometric functions we tested.

Letter identity

Another confound that we wanted to investigate is one that might be relevant in many types of experimental studies where it has been practise to average performance over letter identities; in our study this could have caused the shape of the psychometric function to be corrupted. Therefore we produced enough data so that we could investigate if the form of psychometric function would depend on whether we averaged over the individual letter identities or not. Our results showed that there were significant differences between the parameters of the psychometric functions for the individual letter identities, however our results also showed that, averaging or not, the exponential remained the poorest psychometric function that we tested. Similarly, averaging or not, the squared-logistic and the log-logistic remained the best psychometric functions that we tested.

Confusion matrices

A popular approach to single-letter identification has been to study letter confusion matrices (Townsend, 1971; Gervais, Harvey, & Roberts, 1984; Liu, Klein, Xue, Zhang, & Yu, 2009). From these studies it is evident that each letter identity posses unique confusion characteristics. Similar letters are presumably more confusable, so the general aim in confusion studies is to extract the feature space of letter coding based on the letter's confusability. A recent development is that it has been demonstrated that letter identification requires the processing of a set of features that all seem to be processed at different timescales (Fiset et al., 2008a; Fiset, Blais, Ethier-Majcher, et al., 2008a). Since each letter identity is composed of a unique feature set the later finding seems to suggest that information available for discriminating between letters evolves as a function of exposure duration. Analysing Figure 7 we found evidence that the pattern of confusion changes as a function of exposure duration. Also we saw that similar letters

were more often confused. This later pattern was especially evident for the letter Q. For subject MK it was striking how all roundish letters: B, C, D, G, J, P, Q, R, S and U were the ones that most often were perceived to be an O. Even though we did not attempt to clarify the constituent features of visual letter based on the confusion matrices, we still think that letter confusion matrices could potentially reveal interesting knowledge about the temporal order in which visual features are processed.

Guessing rates

We analysed guessing by calculating prior probabilities for reporting the individual letter identities. We investigated how the prior probabilities depended on both previously presented and previously reported letters. Our analysis showed that both presented and reported letters influence reports 8 letters ahead even though these are presented about 24 seconds later. Comparing across subjects we found that reporting a letter implies an even longer-lasting impact (35-50 seconds) on later reports than do letters presented.

Also we investigated how the prior probabilities developed over experimental sessions. We found systematic variation in the prior probabilities over session, especially this was clear for subject MK who over sessions shifted from reporting A's to reporting K's and O's. Further we suggest the reader to compare Figure 10 with Figure 6, where it is interesting to observe that exactly A, K and O appear as outliers for subject MK. We have no bulletproof explanation why this is, but a thought is that it might be related to some training effect combined with the observed change in prior probabilities over sessions. Comparing between subjects we see that the prior probabilities depend heavily on the individual subject.

Summary

Analysing our data we fitted six different psychometric functions to our own original data as well as original data from (Shibuya & Bundesen, 1988) and (Bundesen & Harms, 1999). As psychometric functions we used: the exponential, the gamma, the Weibull, the ex-Gaussian, the log-logistic and finally the squared-logistic distribution, which we developed ourselves. For our own original data as well as for the original data from (Shibuya & Bundesen, 1988) and (Bundesen & Harms, 1999) we calculated AIC and BIC for each of the six psychometric functions that we tested. Evaluating these measures we noted a unanimous result, namely that the log-logistic and the squared-logistic were the most optimal and the exponential the least optimal of the psychometric functions we had tested. Our results strongly indicate that performance in single, whole and partial report can be modelled more optimally if one either uses the 2-parameter log-logistic or the 3-parameter squared-logistic psychometric function rather than the 2-parameter exponential that has traditionally been assumed in TVA. Our results improve accuracy in TVA paradigms, which count numerous applications with respect to diagnostics of cognitive disorders and certain types of vision deficits. TVA has particularly been applied to the whole and partial report paradigm, but in addition to that our new psychometric function can also be used directly to address more accurately performance in single-object report, which has quite interestingly been shown to be a good indicator of Alzheimer' disease.

7.1 *Summary of contributions*

In this thesis we have provided a number of important contributions:

- It was shown that the 2-parameter log-logistic and the original 3-parameter squared-logistic are better distributions than the exponential (AIC- and BIC-wise) for modelling the psychometric function for single-letter identification as a function of exposure duration. Furthermore, as is also justified in NTVA, we showed that the appropriate distribution has a non-monotonic hazard rate.
- We formulated the generalized FIRM equations which allow the insertion of any desired psychometric function (rather than only the exponential) into TVA. The generalized FIRM equations represent an example of how choosing the most plausible psychometric function does not in any way exclude the usage of TVA. In fact, insertion of an alternative psychometric function enabled TVA to provide an even better description of whole and partial report data. Inserted into the generalized FIRM equations, the log-logistic and the squared-logistic were the psychometric functions tested that enabled the best description of whole and partial report data.
- We constructed a new type of pattern mask to prevent that the mask should bias the individual letters differently. To what extent this purpose was actually served by the mask, we do not know. Our mask algorithm generates masks that are constructed so that specified features of the masks depend on the features of the specific stimulus set used, and so it would seem that the algorithm can be re-used to suit other sets of stimuli if needed.
- We showed that there is a significant difference in the difficulty of reporting the 26 different letter identities. However, averaging across letter identities did not alter the fact that the log-logistic and squared-logistic psychometric functions were the most appropriate ones that we tested.
- We showed that the limited, but varying capacity of the visual short-term memory could be explained as the outcome of a neural mechanism. Such mechanism might be rooted in interactions between competing neuron ensembles, this can be modelled by the LCA model, and so this might give promise to this type of model.
- Finally, as will be presented shortly in Section 7.2, the thesis also suggests five topics for future work. These topics cover diverse themes such as: 1) possible application areas in cognitive diagnostics, 2) neurodynamical modelling of perception, 3) extraction of the visual features representing objects in our brains, 4) psychometric functions and assumptions as well as 5) duration of visual short-term memories.

7.2 Directions for future work

Someone once said that doing research to answer scientific questions is like throwing logs on a fire in order to light up a dark night. In some way the more logs you throw, the more darkness seem to surround you. Our scientific work undoubtedly answered some of the questions we posed; however as seems generally true, our answers also raise an even larger set of new questions to be addressed in the future. Some of the more interesting of these questions are:

- 1) *Could the models that we developed improve diagnostics of Alzheimer's patients, and what about other patient groups, people with word blindness, reading disabilities, people with HD, multiple sclerosis?*

Our work offers a psychometric function that can model accuracy in single-object identification with unprecedented accuracy. Inserting our new psychometric function into TVA we further provide a better means for modelling also whole and partial report. Single, whole as well as partial report has all shown to provide information on the state of Alzheimer's disease (A. D. Baddeley et al., 2001; Bublak et al., 2006; Cronin-Golomb et al., 1995; Mendola et al., 1995). Especially single-object report was shown to be perhaps the most promising cognitive task to indicate AD (Cronin-Golomb et al., 1995; Mendola et al., 1995). We speculate that imposing either the log-logistic or the squared-logistic when fitting single-letter data, is likely to provide an even more efficient predictor of AD. Also if other parameters than the threshold (Mendola et al. used only the 50 % percent threshold) contains information on the deficit then combining the parameters from the psychometric function, for instance the threshold and the slope for the log-logistic, could be a feasible way to build a more accurate estimator of the cognitive deficit. Performance in other patient groups would of course also be interesting to investigate.

- 2) *What type of accumulator models would show a first-passage-time distribution that mimics the hazard function of distributions such as the log-logistic or the squared-logistic? To what extent would the Ornstein-Uhlenbeck show this behaviour?*

We showed that the log-logistic and the squared-logistic were well-suited distributions to use as the psychometric function for visual identification as a function of exposure duration. Further we noted that the psychometric function seems to show a non-monotonic hazard function. If we relate this to the class of stochastic accumulator models that we introduced earlier, then this means that we need only to focus on the subset of these which possess such non-monotonic first-passage-time distribution. As the LCA model can be reduced to the OU model in the case of single elements being presented, we therefore speculate to what extent the OU or similar types of models would show this behaviour (Aalen et al., 2008; Aalen & Gjessing, 2001, 2004). This might be relevant to investigate in future studies.

- 3) *What would be an optimal way to analyse the temporal development observed in the letter confusion matrices, and does confusability tell us something about letter features and the binding of these?*

We think that datasets such as ours might be used to tell us something about the order in which features are processed and combined in order to allow for identification. As an example we saw that our dataset could be used to track the temporal time-course of letter confusability. In Figure 7 we showed the log posterior mean of reporting an O as a function of exposure duration and identity of the letter presented. It was clear that for roundish letters (such as for Q) the log posterior mean of reporting an O seemed to first increase and then to decrease as a function of exposure duration. If the reported letter depends on the features processed then this might suggest that some features are processed earlier on than others (Fiset, Blais, Arguin, et al., 2008b; Fiset et al., 2008b). Assuming that different letter features are processed at different time-scales, then a possible way to go would be to construct a clever factor analysis type of algorithm capable of extracting estimates of the visual features used for human letter identification. These estimates might then be based on letter confusion matrices such as ours.

- 4) *Is it possible to devise a unified psychometric function for visual identification that takes into account duration, contrast and size? Furthermore, would such a psychometric function satisfy the laws of Ricco and Bloch?*

For the detection task Bloch's law predicts how performance varies as a function of time and contrast. It is not clear however, whether this law also holds in the case of identification (Smith, 1998; Palmer, Huk, & Shadlen, 2005; Scharnowski, Hermens, & Herzog, 2007), for example in single-letter report. Therefore it would be relevant to investigate this. A law similar to Bloch's law is Ricco's law (Schwartz, 1998; Strasburger, 2005). Ricco's law defines how performance in a detection task varies as a function of contrast and size; obviously one could also investigate how the psychometric function depends on the size of the stimulus, and whether this dependence is in accordance with Ricco's law.

- 5) *Does guessing depend on previously reported - or previously presented letters and should we perhaps expect guessing rates to vary over sessions?*

Our work showed that guessing rates were significantly influenced by both previously reported and previously presented letters. The duration of the effect from presented letters seemed to be at least 24 seconds while the effect from reported letters was even longer (35-50 seconds). We should note that our investigation was not corrected for the cases in which a letter in a previous trial was *both* presented and reported, though it would seem that it would be more likely to report a letter if it was presented. Therefore a direction for future work could be to perform such kind of correction, so that the effect of presented and reported letters on later reports can be statistically separated. If the duration and strength of the effect of previous presented and reported letters on later reports can be described, then this information could be used to create a more precise model for deciding the probability of a correct report. Also, related to the observation that guessing rates seem to develop systematically over sessions, it might be interesting

to study why this is. Perhaps the subject shifts from attending an initial set of features to gradually attending a feature set that more effectively provides a means to discriminate between the various stimulus categories.

Appendix A Paper for ICCM2009

Towards a neural network model of the visual short-term memory

Anders Petersen¹ (ap@imm.dtu.dk)
Søren Kyllingsbæk^{1,2} (sk@imm.dtu.dk)
Lars Kai Hansen¹ (lkh@imm.dtu.dk)

¹Department of Informatics, Technical University of Denmark,
Richard Petersens Plads, B. 321,
2800 Kgs. Lyngby, Denmark

²Department of Psychology, Copenhagen University
Linnésgade 22,
1361 Copenhagen, K, Denmark

Abstract

In this paper a neural network model of visual short-term memory (VSTM) is presented. The model aims at integrating a winners-take-all type of neural network (Usher & Cohen, 1999) with Bundesen's (1990) well-established mathematical theory of visual attention. We evaluate the model's ability to fit experimental data from a classical whole and partial report study. Previous statistic models have successfully assessed the spatial distribution of visual attention; our neural network meets this standard and offers a neural interpretation of how objects are consolidated in VSTM at the same time. We hope that in the future, the model will be developed to fit temporally dependent phenomena like the attentional blink effect, lag-1 sparing, and attentional dwell-time.

Keywords: visual attention, visual short-term memory, the magical number 4, winners-take-all network

Introduction

For everyday life, it is important for us to be able to perceive, comprehend, and react to events in our environment. Often, our rate of success is heavily dependent upon how efficient and how fast we can process, interpret and react to sensory stimuli, e.g. like when we are driving a car.

In the following we shall refer to *visual attention* as the process that enables us to focus our processing resources to certain important objects in the visual scene. Following the theory of visual attention (TVA, Bundesen, 1990) we assume that features have already been extracted and objects successfully segregated on the basis of their individual feature spaces. Our model deals with the important question of how only a limited sub span of all objects are actually selected and further encoded into VSTM.

Cattell already in the late 19th century demonstrated a surprising limit in how many objects that can be perceived at the same time – a limit only about 4 objects which may be held in the VSTM at the same time (Cattell, 1886; Cowan, 2000). This finding is independent of the number of objects visually presented at the same time (Sperling, 1960). Evidence further exist that the “magical number” of 3-to-4 objects is largely independent of how many features are

encoded for each object, i.e. the complexity of the visual object, does not hold an influence on the memorial capacity of the VSTM; see (Luck & Vogel, 1997), but see also (Alvarez & Cavanagh, 2004).

Modelling the function of the VSTM, it is essential that the inherent capacity limitation is properly mimicked, since it seems a fundamental limit of the system. Most likely the VSTM would be heavily overloaded, should the system lack the ability to represent only the most salient of the visually appearing objects

The model

The model that we are presenting in this paper can actually be understood as three consecutive processes (See Figure 1).

The first process is simply extraction of visual features, we speak of this process as '*object matching*', since we find it relevant to think that objects in the visual field are to some extent 'matched' against objects representations in Visual Long-Term Memory (VLTM). In this paper we do not consider the problem of which feature extraction techniques are biologically most plausible or perhaps technically most appropriate to use.

The second process that we shall consider in more detail is '*the attentional race*'. According to Shibuya & Bundesen (1988), all objects in the visual scene take a place in what one could think of as a race to become encoded. In Shibuya & Bundesen's race model, the 'odds' that a given object is selected as a winner in the race is directly related to the rate value with which the object participates. It is worth noting that the race is a stochastic, rather than a deterministic process, meaning that no one can beforehand predict readily which objects will win the race.

The third and last process that we shall consider is that of '*storage*' of object representation in VSTM. Inspired by (Usher & Cohen, 1999) we propose a competitive neural network model of VSTM, directly linking with several important assumptions expressed in Bundesen's Theory of Visual Attention (Bundesen, 1990).

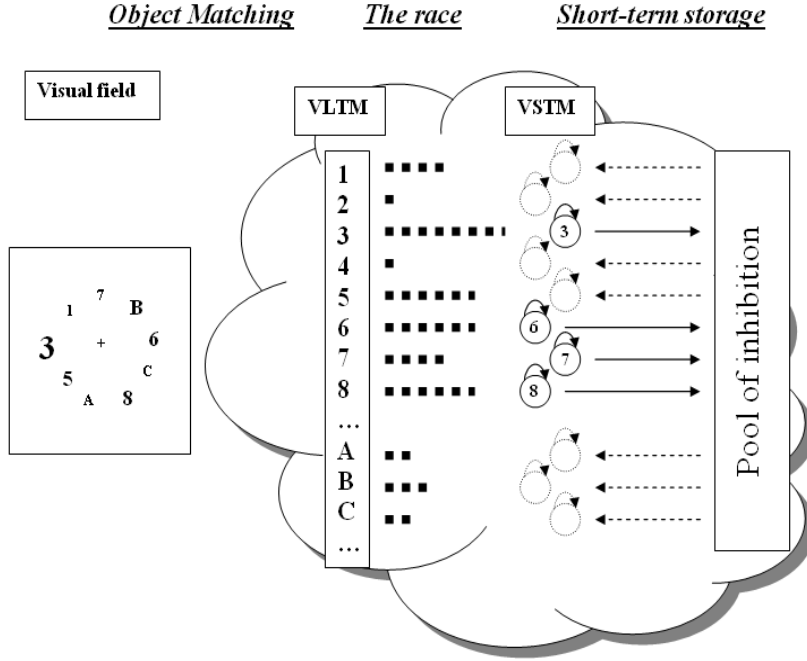


Figure 1: The Model Scheme – a partial report example. The task is to report the targets, i.e. digits and ignore the distractors, i.e. letters. The model predicts how visual elements participate in a race, where the winners become selected to be encoded in visual-short-term memory. Generally targets are processed faster than distractors, however we also see that in the example *homogeneity* is not assured, i.e. the targets (and distractors) are not of equal size (could also be contrast, letter type etc.) and therefore in the example they are illustrated as being processed with slightly different rates.

The neural theory of visual attention

The theory of visual attention (TVA) proposed by Bundesen (1990) is a unified theory of visual recognition and attentional selection. TVA provides a mathematical framework describing how the visual system is able to select individual objects in the visual field S , based on the visual evidence, η and the setting of two different types of visual preference parameters (pertinence, π and bias, β), representing the influence from higher cortical areas, including VLTM.

The output of the TVA-model is a set of rate parameters v that are directly related to the probability that a given characterization, *object x belongs to category i* , is encoded into the VSTM. The rate parameters are given by:

$$v(x, i) = \eta(x, i) \beta_i \frac{w_x}{\sum_{z \in S} w_z} \quad (1)$$

Where the attentional w_x weight of object x is:

$$w_x = \sum_{j \in R} \eta(x, j) \pi_j \quad (2)$$

Here $\eta(x, i)$ is defined as the strength of the sensory evidence that object x belongs to the visual category i . The pertinence of the visual category j is denoted by π_j and setting of these values effectively implements the so-called filtering mechanism. The perceptual decision bias of a visual category i is denoted by β_i and setting of these values conversely implements a complementary mechanism called pigeonholing.

The filtering mechanism increases the likelihood that elements belonging to a target category are perceived, without biasing perception in favor of perceiving the elements as belonging to any particular category.

Pigeonholing, conversely changes the probability that a particular category i is selected without affecting the conditional probability that element x is selected given that category i is selected.

A neural interpretation of TVA is given in (NTVA, Bundesen, Habekost, & Kyllingsbæk, 2005). Basically here pigeonholing (selection of features) is considered an increase in the rate of firing of neurons while filtering (selection of objects) is considered an increased mobilization of neurons.

Corresponding to the interpretation in NTVA the fraction $w_x / \sum w_z$ in equation (1), which is the relative attentional weight of object x compared to the weight of all objects z in the visual field S , can be directly interpreted as the relative fraction of neurons allocated to process a given object x ,

compared to the total number of neurons processing just any object z belonging to the visual field S .

Each and every encoding generally takes the form *object x belongs to category i* .

Denoting the set of all features as R the total processing capacity, can be considered a constant C , which equals the sum of all encoding rates v ; see (Bundesen, 1990).

$$C = \sum_{x \in S} \sum_{i \in R} v(x, i) \quad (3)$$

Shibuya and Bundesen (1988) assume target as well as distractor homogeneity in their whole and partial report paradigm. This means that processing capacity is distributed equally among targets as well as among distractors. When this is the case the rates of encoding for targets, v_T and for distractors, v_D can be calculated according to the formulas:

$$v_T = \frac{C}{T + \alpha D} \quad v_D = \frac{\alpha C}{T + \alpha D} = \alpha v_T \quad (4)$$

Where T and D denote the number of targets and distractors presented, respectively. The ratio of discrimination between distractors and targets is denoted α .

The effective exposure duration τ is smaller than the actual exposure duration t by an amount t_0 corresponding to the temporal threshold before conscious processing begins. However the effective exposure duration can not be negative so computationally it is set to:

$$\tau = \max(0, t - t_0) \quad (5)$$

In the neural network model that we shall now describe we adopt the parameters C , α and t_0 and further, following Bundesen, we make use of equation (4) and equation (5).

The neural network model of VSTM

In TVA object features are encoded independently, and further there is the assumption that only one feature needs to be encoded for the object to be stored in VSTM. On the other hand; and in agreement with (Luck & Vogel, 1997), several features of the same object can be in the encoded state, and still it will only count as if one object is stored in VSTM. For this reason, and because here we are concerned about objects rather than features encoded, we simply sum over the entire number of object features, and in this way we obtain the total encoding rate v_x for object x :

$$v_x = \sum_{i \in R} v(x, i) \quad (6)$$

An object x can enter VSTM once it receives external excitation, G taking the shape of Poisson distributed spike trains, arriving with the rate parameter v_x . (See Figure 2).

A neural assembly that has obtained a positive level of activation will automatically seek to re-excite itself, so that it can stay in VSTM, at the same time trying to inhibit activation in other neuron assemblies representing other objects, i.e. working to suppress other object from co-temporally being stored in VSTM.

The initial condition for the simulations is that all neuron assemblies start with an activation of zero, i.e. no objects are initially stored in VSTM. As a consequence neither re-excitation nor lateral inhibition exists, before the assemblies are externally activated.

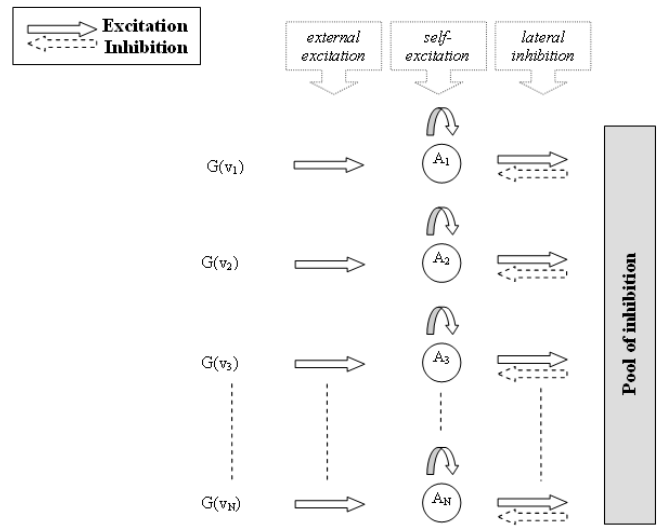


Figure 2: The neural network model of VSTM. The total number of neuron assemblies is N and each assembly is represented by a level of activation A

Implementation

The activation A_x of neuron assembly x (representing object x) is given by the first order differential equation:

$$\frac{dA_x}{dt} = -A_x + \alpha^* F(A_x) - \beta^* \sum_{z \neq x} F(A_z) + \gamma^* G(v_x) \quad (7)$$

The above equation characterizes a leaky accumulator model. There is passive decay of the activation towards the rest level, with a time constant chosen as 1, reflecting the time scale that physiologically is observed with synaptic currents (Usher & Cohen, 1999).

F is a squashing function that keeps the activation within bounds:

$$\begin{aligned}
F(A) &= 0, & \text{for } A \leq 0 \\
F(A) &= \frac{A}{1+A}, & \text{for } A > 0
\end{aligned} \tag{8}$$

As a consequence of the squashing function F , the parameter α^* is the limiting value of maximal self-excitation that assemblies can up-hold and the parameter β^* is the limiting maximal value of inhibition that can be sent from one assembly to another.

Also the model assumes we can not have negative self-excitation, i.e. self-inhibition and further the model does not implement any terms that could account for excitation laterally between the assemblies. The latter effect could for instance be included if one wanted to account for semantically related objects and their effect on the number of reported objects.

The attentional significance that object i is present in the visual field R is represented by the encoding rate v_i . In our model we follow the approach from (Bundesen, 1990) and interpret this rate as the firing rate of a Poisson spike generator G . Hence γ^* characterizes the amplitude of the Poisson distributed input spikes arriving to the neuron assembly x .

The model was implemented in Matlab's Simulink toolbox. At least in the operated parameter domain we judge the stiffness of the system to be negligible so for simplicity we numerically apply Euler integration¹.

Model performance

The dataset

The data covers the performance of a single subject, participating in an extensive series of whole and partial report experiments. The subject was instructed to report targets, i.e. digits while ignoring distractors, i.e. letters displayed on an imaginary circle around a small fixation cross at the center of the screen. In practice experimental trials covered twelve whole and partial report conditions. In these the number of targets, T was between 2 and 6 and the number of distractors, D was between 0 and 6. Further, exposure durations t were varied systematically at 10, 20, 30, 40, 50, 70, 100, 150 and 200 ms. Each experimental condition was repeated 60 times but trials were mixed so that the subject had no a-priori knowledge of the experimental condition. Moreover trials were grouped into blocks to minimize the element of fatigue. Each presented character was immediately followed by a mask lasting for 500 ms. Further information can be found in (Shibuya & Bundesen, 1988).

¹ Assuming that only one spike should be allowed in each time step we must keep the integration step size sufficiently small. If the processing capacity C is 60 Hz, and the integration step size is kept at $dt = 0.001$, then the risk that two or more spikes will be present in a given time step is as low as 0.36 %.

Performance of the neural network model

Figure 3 shows accumulated score distributions. The score is defined as the number of targets reported correctly. The upper most curve represents the accumulated score of $j = 1$, i.e. the probability of reporting 1 or more targets correctly. Other curves represent accumulated probabilities for reporting at least 2, 3, 4 or even 5 targets.

Shibuya and Bundesen (1988) proposed a mixture model, mixing probabilities obtained with using a statistical model that assumed memorial capacities of either $K = 3$ or $K = 4$ respectively.

There is a relatively close fit between the proposed mixture model and the empirical data. We see however that data points obtained with exposure duration around 50 ms are generally under-fitted and more noticeably the model does not account for cases where more than 4 targets are reported, as is actually the case in two out of three of the lower most plots.

What we observe with the previous model can be considered a trade-off between two conflicting demands. The first demand is to fit the initial part of the curves, i.e. the larger the processing capacity C the steeper the curves will rise, on the other hand the second demand, which is to keep the score distribution reasonably low for long exposure durations, require that the processing capacity C is not set too high. Hence the setting of C is set subject to a compromise.

Addressing the performance of our neural network model we think it clearly meets the standard of Shibuya and Bundesen's model. The neural model does however seem to have some trouble predicting 4 recognized items in the situations where no distractors were presented. Possibly this misfit can be diminished by running a more exhaustive optimization of model parameters. The parameters used for producing the figure were: $\alpha^* = 5$, $\beta^* = 0.1$, $\gamma^* = 2$, $C = 61.5$ Hz, $t_0 = 23$ ms and $\alpha = 0.367$. Moreover, and in contrast to Shibuya and Bundesen's model, our new model readily demonstrates its capability of predicting extreme cases, where more than 4 objects are reported.

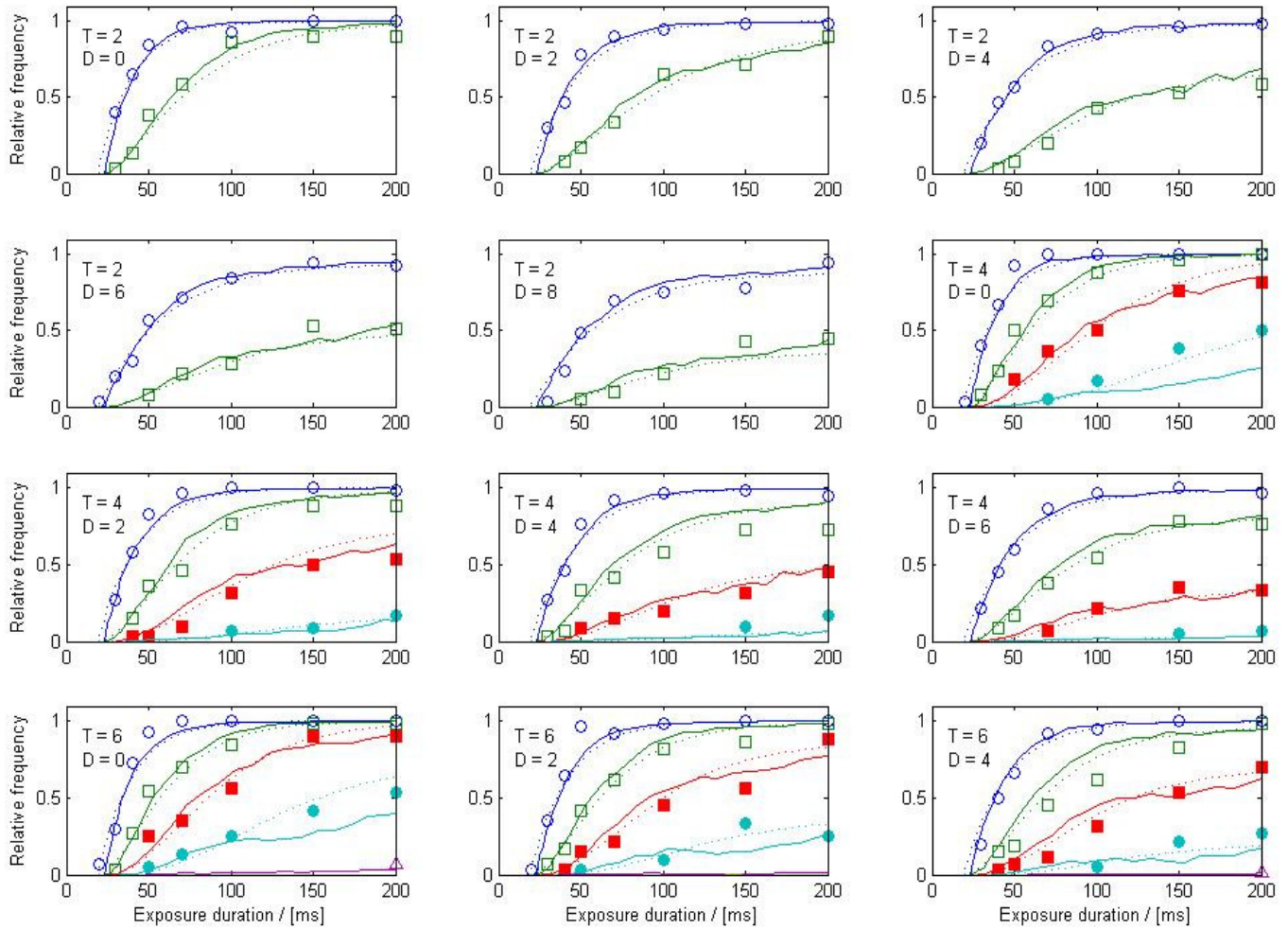


Figure 3: Accumulated score distribution for subject MP in (Shibuya & Bundesen, 1988). Probability of correctly reporting at least 1 target (blue, open circles), 2 targets (green, open squares), 3 targets (red, closed squares), 4 targets (cyan, closed circles) and 5 targets (magenta, open triangles). Empirically found values are plotted with symbols as markers. The dotted lines represent the fit by Shibuya & Bundesen (1988). Solid lines represent the performance of our neural network model. T and D denote the number of targets and distractors presented, respectively.

Discussion

This work represents an attempt to integrate the Theory of Visual Attention (Bundesen, 1990) with a simple type of winners-take-all type of network (Usher & Cohen, 1999), in the sense that the later implements a limited storage capacity of VSTM. Our new dynamic model of visual attention and VSTM is able to account for the complete set of data from whole and partial report experiments. Where the previous account by Shibuya and Bundesen (1988) treated extreme scores as outliers, the new model encompasses these as natural consequences of the internal

dynamics. Further, the model explains VSTM capacity and consolidation as the result of a dynamic process rather than as a static store, which capacity is independent of processing capacity and the attentional set of the subject.

From daily life we know that humans are able to identify a very larger number of different objects. Therefore, we might think that we would have to include a neural assembly for each of these many objects candidates in our model of identification. However, what we shall argue is that our model's predictions are not affected if irrelevant neural assemblies (representing non-stimuli type of objects) are not included in the model, a useful feature which we of course make use of when we simulate with the model. The reason for this is that in the model only activated neural

assemblies affect other assemblies, and so there is no lateral inhibition from inactive neural assemblies (which irrelevant assemblies tend to be) upon any other assembly. This means that adding more irrelevant assemblies generally does not affect our conclusions, except that computationally simulations become slower.

The model described gives no account of identification of individual features of an object; however it would be possible to approach this situation by having one neural assembly in the network per object feature, rather than just one neural assembly per object. In this case assemblies representing features that belonged to the same object might be modeled as having little or no lateral inhibition, ensuring that several features of the same object can be encoded without taking up additional VSTM storage space (Luck & Vogel, 1997).

Speaking of adding more neural assemblies, we ought to touch upon what it is that we think an assembly represents. Does the assembly manifest itself in one or more neurons, and how would this relate to efficient or distributed processing? The way we think about the model is that the assemblies conceptually represent different states of neural activation. As assumed, these states interact and as we have described we suppose that feedback mechanisms play an important role in keeping the activation of the assembly sustained, allowing for visual short-term memories.

A possible confound of the model is that it does not consider internal noise, which is likely to play a key role in many neural systems. A way to deal with this would be to transform the input stage (the Poisson distributed spike trains, arriving with the rate parameter ν) to a stochastic diffusion process with Wiener noise process included. For this to make sense the activation threshold for consciousness would have to take a higher value than the level of initial activation.

In future studies, we think it would be relevant to explore the implication of transforming the model into a stochastic differential equation as mentioned above. Because the model is temporally dependent it would also be interesting to know if it would be able to address the dynamic consolidation in VSTM found in temporally extended paradigms such as the attentional blink paradigm and studies of attentional dwell time; e.g. (Ward, Duncan, & Shapiro, 1996). Here, consolidation in VSTM is strongly dependent on competition between items already encoded into VSTM and visual items presented at a later point in time. Incorporation of such a competitive process follows naturally from the dynamic architecture of the present model.

Acknowledgments

We would like to thank Hitomi Shibuya & Claus Bundesen for access to the experimental data in (Shibuya & Bundesen, 1988). Also we would like to thank Marius Usher for kindly verifying parameters and settings in (Usher & Cohen, 1999). Further we would like to thank the anonymous reviewers for providing helpful suggestions as well as relevant comments.

References

- Alvarez, G., & 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.
- Bundesen, C. (1990). A Theory of Visual Attention. *Psychological Review*, 97(4), 523-547.
- Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2005). A Neural Theory of Visual Attention: Bridging Cognition and Neurophysiology. *Psychological Review*, 112(2), 291-328.
- Cattell, J. M. (1886). The inertia of the eye and brain. *Brain*, 8, 295-312.
- Cowan, N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24(1), 87-185.
- Luck, S., & Vogel, E. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657), 279-281.
- Shibuya, H., & Bundesen, C. (1988). Visual Selection from Multielement Displays: Measuring and Modeling Effects of Exposure Duration. *Journal of Experimental Psychology Human Perception Performance*, 14(4), 591-600.
- Sperling, G. (1960). The Information Available in Brief Visual Presentations. *Psychological Monographs*, 74(11), 1-29.
- Usher, M., & Cohen, J. (1999). Short Term Memory and Selection Processes in a Frontal-Lobe Model. In *Connectionist Models in Cognitive Neuroscience* (pp. 78-91). Birmingham: Springer-Verlag.
- Ward, R., Duncan, J., & Shapiro, K. (1996). The slow time-course of visual attention. *Cognitive Psychology*, 30(1), 79-109.

Appendix B Paper submitted

The effect of exposure duration on visual character identification in single, whole and partial report

Keywords: psychometric function, character identification, Theory of Visual Attention, visual
short-term memory, exposure duration

Abstract

The psychometric function of single letter identification is typically described as a function of stimulus intensity. However, the effect of stimulus exposure duration on letter identification remains poorly described. This is surprising because the effect of exposure duration has played a central role in modelling performance in whole and partial report tasks in which multiple simultaneously presented letters are to be reported (Shibuya & Bundesen, 1988). Therefore, we investigated visual letter identification as a function of exposure duration. On each trial, a single randomly chosen letter (A-Z) was presented at the centre of the screen. Exposure duration was varied from 5 to 210 milliseconds. The letter was followed by a pattern mask. Three subjects each completed 54,080 trials in a 26-Alternative Forced Choice procedure. We compared the exponential, the gamma and the Weibull psychometric functions, all of these having a temporal offset included, as well as the ex-Gaussian, the log-logistic and finally the squared-logistic, which is a psychometric function which we believe, has not been described before. The log-logistic and the squared-logistic psychometric function fit well to experimental data in both the present study and in a previous study of single-letter identification accuracy. Also, we conducted an experiment to test the ability of the psychometric functions to fit single-letter identification data, at different stimulus contrast levels; also here the same psychometric function prevailed. Finally, after insertion into Bundesen's Theory of Visual Attention (Bundesen, 1990), the same psychometric functions enable closer fits to data from a previous whole and partial report experiment.

Introduction

A visual scene typically contains several objects, one or more which are of special importance for us to identify and some that are not. Being able to quantify and model the accuracy of visual object identification in multi-object displays is important for such diverse areas such as reading speed and learning (Pelli & Tillman, 2007; Rasinski, 2000), traffic safety (Baldock, Mathias, McLean, & Berndt, 2007; Richardson & Marottoli, 2003), human-computer interaction (Chen & Chien, 2005; Chien, Chen, & Wei, 2008) and diagnostics of perceptive and cognitive disorders (Behrmann, Nelson, & Sekuler, 1998; Cheong, Legge, Lawrence, Cheung, & Ruff, 2007; Habekost & Starrfelt, 2008).

Whole and partial report experiments concern visual identification in multi-element displays which can be thought of as simplified visual scenes compared to the more complex ones that we typically encounter in real life. In *whole report* a number of simultaneously presented target elements are to be reported by the subject. *Partial report* is similar to whole report, except that distractor elements (elements that are not to be reported) are concurrently included in the display.

Bundesen's Theory of Visual Attention (TVA, Bundesen, 1990) offers a quantitative model linking perception of single isolated objects to perception of multiple objects in whole and partial report experiments. The theory assumes that the total amount of perceptual processing resources, which determine the rate of perceptual processing, is limited. The processing resources are distributed evenly among target objects. Through attentional filtering a proportionally smaller amount of processing resources is allocated to distractor objects. After processing resources are allocated, all objects participate in a race for being encoded into the visual short term memory (VSTM), which

has a limited storage capacity. Since the time course of encoding is so central to TVA, the theory has mainly been applied to experiments in which stimulus exposure duration has been the independent variable. In Appendix B we present relevant formulas from TVA. When applied to single objects TVA is reduced to the psychometric function for object identification as a function of stimulus exposure duration but, surprisingly, few studies (for an exception see Bundesen & Harms, 1999) have studied this topic. This is our motivation for studying the psychometric function for letter identification as a function of stimulus exposure duration in more detail.

A psychometric function $\psi(t; \theta, \gamma, \lambda)$ quantifies the probability of a correct report as a function of some stimulus attribute t , which in our case is exposure duration. It is characterized by a number of parameters that include the parameter set θ of the function F as well as two additional parameters, γ and λ , that denote the guessing and lapsing probabilities, respectively. We define the guessing probabilities as the fraction of times an un-informed observer presses (intentionally or accidentally) each of the keys included in the response set. The lapsing probability we define as the relative fraction of accidental key presses, averaged over all keys in the response set. The psychometric function ψ , which includes correction for guessing and lapsing, can be written as

$$\psi(t; \theta, \gamma, \lambda) = F(t; \theta) \cdot (1 - \lambda) + (1 - F(t; \theta)) \cdot \gamma = \gamma + (1 - \gamma - \lambda) \cdot F(t; \theta)$$

What we shall generally speak of as the *psychometric function* is the function F , i.e. the psychometric function *after* correction for guessing and lapsing (Treutwein & Strasburger, 1999; Wichmann & Hill, 2001). In the following we will describe six psychometric functions, which, for various reasons, are plausible candidates for describing letter identification as a function of exposure duration.

The exponential distribution with a temporal offset included was used as the psychometric function in TVA (Bundesen, 1990). Ignoring the temporal offset, this psychometric function assumes that encoding into VSTM can be considered events from a *homogenous* Poisson process, for which the waiting time is well known to be exponentially distributed. The exponential distribution has the parameter set $\theta = \{\nu, \mu\}$, where $\nu > 0$ is the rate and $\mu > 0$ is the temporal offset of the Poisson process, and the distribution is defined by:

$$F(t; \theta) = 1 - e^{-\nu(t-\mu)} \quad \text{for } t \geq \mu \quad \text{and} \quad F(t; \theta) = 0 \quad \text{for } t < \mu$$

From a psychophysical perspective, it is strange that there should be a fixed temporal offset before encoding can take place, and that after this point the encoding rate is constant; instead, we find it more plausible that the encoding rate rises as a smooth function of exposure duration. The gamma, the Weibull and the ex-Gaussian distributions all represent generalisations of the exponential distribution, and all of these are smooth functions.

The Weibull distribution has the parameter set $\theta = \{\mu, \sigma, k\}$, where $\mu, \sigma, k > 0$. It reduces to the exponential distribution when the shape parameter $k = 1$. The Weibull distribution has previously been used for modelling the psychometric function of visual contrast detection, visual discrimination as well as visual identification when these were investigated as a function of stimulus contrast (Pelli, 1985, 1987; Pelli, Burns, Farell, & Moore-Page, 2006). When the Weibull distribution has $\mu > 0$ it includes an offset. The three-parameter Weibull distribution function is defined as:

$$F(t; \theta) = 1 - e^{-\left(\frac{t-\mu}{\sigma}\right)^k} \quad \text{for } t \geq \mu \quad \text{and} \quad F(t; \theta) = 0 \quad \text{for } t < \mu$$

The gamma distribution has the parameter set $\theta = \{\mu, \sigma, k\}$, where $\mu, \sigma, k > 0$. The gamma distribution corresponds to the waiting-time distribution, when waiting for k independent, and identically distributed events, that each has an exponentially distributed waiting time-distribution and it thus reduces to the exponential distribution when $k = 1$. If correct identification depends on the firing of several independent neural units firing as Poisson processes, then the gamma distribution could describe the psychometric function of identification as a function of stimulus duration. Based on a similar argument the gamma distribution has been fitted to response time distributions (Van Breukelen, 1995; Luce, 1991). Noting that Γ is the complete gamma function, and γ is the lower incomplete gamma function, the three-parameter gamma distribution function is defined as:

$$F(t; \theta) = \frac{\gamma\left(k, \frac{t-\mu}{\sigma}\right)}{\Gamma(k)} \quad \text{for } t \geq \mu \quad \text{and} \quad F(t; \theta) = 0 \quad \text{for } t < \mu$$

The Ex-Gaussian distribution has the parameter set $\theta = \{\mu, \sigma, \tau\}$, where $\mu, \sigma, \tau > 0$. The ex-Gaussian approaches the exponential distribution in the limit, to be exact when $\mu=0$ and $\sigma \rightarrow 0$. The Ex-Gaussian distribution characterises the sum of an exponential distributed variable and a Gaussian distributed variable. Thus, if Gaussian noise is added to the temporal offset, μ , in the exponential distribution the waiting times for perceptual processing would be distributed according to the ex-Gaussian distribution. The ex-Gaussian has been used for modelling reaction-time data (Luce, 1991). Noting that Φ is the Gaussian distribution function, the ex-Gaussian distribution function is defined as:

$$F(t; \theta) = \Phi\left(\frac{t - \mu}{\sigma}\right) - \Phi\left(\frac{t - \mu - \sigma^2 / \tau}{\sigma}\right) \cdot \exp\left(-\frac{t}{\tau} + \frac{\mu}{\tau} + \frac{\sigma^2}{2\tau^2}\right)$$

An important question is what causes the shape of a psychometric function? Clearly the shape must reflect the construct and limitations of the physical mechanism underlying perception, i.e. it must reflect the neural activity in the task relevant areas of the brain. It has previously been demonstrated that individual sensory neurons show response functions (firing rate vs. stimulus intensity) that closely resemble the psychometric function seen in detection tasks, such as the logistic distribution (Lansky, Pokora, & Rospars, 2007). In NTVA, which is a neural interpretation of TVA (Bundesen, Habekost, & Kyllingsbæk, 2005), it is assumed that the rate at which stimuli are perceptually processed is proportional to neural firing rates in the visual cortex. Mathematically, the processing rate is the *hazard* rate. Further descriptions of the concept of hazard rate can be found in (Luce, 1991; Van Zandt, 2002; Aalen & Gjessing, 2001). The processing rate is explicit in the exponential function where it equals the parameter, ν . For other psychometric functions the hazard rate is generally not explicit but can easily be derived (see Appendix B).

Exposing cats and monkeys to transient stationary gratings with a duration of 200 ms, (Albrecht, Geisler, Frazor, & Crane, 2002) mapped out the instantaneous firing rates of responsive neurons in the visual striate cortex. The typical temporal profile of the firing rates is similar to the profile that (Bundesen & Habekost, 2008, p. 116) expect follows the abrupt onset of a stimulus: ‘*When a stimulus appears abruptly (a kind of successive contrast), firing rates of typical neurons responding to the stimulus first increase rapidly, then reach a maximum, and finally decline and approach a somewhat lower, steady state level*’. Therefore, it is likely that the psychometric function for object

identification as a function of exposure duration has a non-monotonic hazard rate. Accordingly, a preliminary report (Shibuya, 1994) described the hazard function in a 2-AFC discrimination task, in which exposure duration was varied, as having a non-monotonic hazard rate. However, all of the functions described above have monotonic hazard rates. Therefore, we find it worthwhile to consider also two psychometric functions that have similar temporal profiles, i.e. non-monotonic hazard functions.

The log-logistic is such a distribution, since with appropriately chosen parameters; it has a uni-modal, and hence non-monotonic hazard function. The Log-logistic (or Fisk) distribution is the probability distribution of a random variable whose logarithm has a logistic distribution. It has been used for modelling various kinds of diffusion processes (Brüederl & Diekmann, 1995; Diekmann, 1992). Also it has been used for modelling proportion correct in single-digit identification as a function of contrast (Strasburger, 2001). The Log-logistic distribution has the parameter set $\theta = \{\mu, \sigma\}$, where $\mu, \sigma > 0$. Noting that the parameter σ determines the steepness, and μ is the median survival time, the log-logistic distribution function is defined as:

$$F(t; \theta) = \frac{1}{1 + \left(\frac{t}{\mu}\right)^{-\sigma}}$$

The squared-logistic distribution is another distribution with a non-monotonic hazard function. We describe the squared-logistic because we found it to be a simple function which has a hazard function that closely resembles the instantaneous firing rate of single neurons in the visual cortex like those depicted in (Albrecht et al., 2002). Compared to the hazard function of the log-logistic distribution the hazard function of the squared-logistic distribution seems to drop off faster after the

peak, and furthermore the hazard approaches the quasi-stationary level V rather than continuing to drop off as $t \rightarrow \infty$. The squared-logistic has the parameter set $\theta = \{V, \mu, \sigma\}$, where $V, \mu, \sigma > 0$. We define the squared-logistic distribution function as:

$$F(t; \theta) = 1 - e^{-\frac{V \cdot t}{\left(1 + e^{\left(\frac{t - \mu}{\sigma}\right)^2}\right)^2}}$$

We named the distribution the squared-logistic because, the shape of the mean cumulative hazard function in the interval between 0 and t as a function of time has the shape of a logistic distribution function squared. This can be seen by dividing the negative exponent (the cumulative hazard function) of the distribution by the size t of the temporal interval. Note that V scales the hazard rate and that it is straightforward to derive the probability density function if needed.

To find the most appropriate psychometric function we evaluate each of the six functions described above on four data sets, two of which are from original experiments and two of which stem from previous experiments.

In Experiment 1 we investigate the psychometric function of single-letter identification as a function of exposure duration. As an initial approach we average performance across the different letter identities although there is no a priori reason to assume that the psychometric function preserves its shape when averaged across several stimuli in an identification task. In fact, the psychometric function for identification of individual letters as a function of contrast was investigated in (Alexander, Xie, & Derlacki, 1997). This study demonstrated that the psychometric function for identification of 10 Sloan letters depended significantly on letter identity. Though we

vary exposure duration rather than contrast, we think it is reasonable to assume that averaging over letter identities may affect the shape of the psychometric function. Therefore, we also fit the 6 psychometric functions to the data without averaging across letter identity.

Experiment 2 is similar to Experiment 1 but the contrast is varied between 11 different levels; the number of repetitions is lowered and shorter exposure durations are used for higher stimulus contrast levels. Bundesen and Harms (1999) showed that the psychometric function for letter identification as a function of exposure duration is very fast for high contrast stimuli. This is problematic because it is difficult to present letter stimuli with such high temporal resolution. Therefore, in Experiment 1 we decided to use a lower contrast level than that used by Bundesen and Harms (1999). The effect of changing the stimulus contrast level was demonstrated by Di Lollo and co-workers (Di Lollo, von Mühlenen, Enns, & Bridgeman, 2004, fig. 4) when they investigated visual single-object identification, also as a function of exposure duration. In their study, proportion correct as a function of exposure duration was plotted for four different contrast levels, and from visual inspection, it is evident that identification accuracy drops as stimulus contrast is decreased; however they did not fit psychometric functions to their data. A reasonable question to ask is, if different stimulus contrast levels might also favour different types of psychometric functions for single-letter identification as a function of exposure duration, or alternatively that the same psychometric function can be used, only with different parameter values. The question is particularly relevant here because the psychometric function determined in Experiment 1 accounts for identification of stimuli that are displayed at a relatively lower contrast than has traditionally been the case in TVA studies (Shibuya & Bundesen, 1988; Bundesen & Harms, 1999).

Bundesen & Harms (1999) investigated the psychometric function of letter identification as a function of exposure duration. In this study the exponential psychometric function was used to model the data from the three subjects each carrying out a total of 4000 trials. Bundesen and Harms (1990) did not however, fit any other psychometric function to their data. Here, we compare the fits of the 6 psychometric functions described above to their original data.

Shibuya and Bundesen (1988) conducted a whole and partial report experiment with two observers each completing 6480 trials. In whole report, observers were presented with 2-6 visual targets (digits). In partial report, up to 8 distractors (letters) were presented with the targets. Observers' performance was recorded as the proportion of scores of j or more (correctly reported targets). Shibuya and Bundesen showed that TVA could account very well for the observed data. Until now TVA has assumed an exponential psychometric function but in Appendix B we show how to generalize it to allow for other psychometric functions by letting the encoding process be a non-homogenous Poisson process. Here, we insert each of the six psychometric functions described above into TVA and test each of these six models against Shibuya and Bundesen's original data.

In summary, our aim is to investigate the psychometric function of visual identification as a function of exposure duration. We evaluate six selected psychometric functions for single letter identification when averaging data over letter identities and also for each individual letter identity. By inserting each psychometric function into TVA we also evaluate their appropriateness for describing performance in whole and partial report.

Methods

Paradigm

The task was to report a single stimulus letter cued by a fixation point and terminated by a mask. The report was carried out as a forced choice procedure with 26 alternatives (26-AFC). A trial commenced with the fixation point marker being displayed for 1000 ms. Immediately after this the stimulus letter was shown. A randomized mask, lasting for 500 ms, followed the stimulus and then the report display consisting of the 26 letters of the alphabet was shown. The subject had to report which stimulus letter he thought was the one presented by typing the letter on a standard Danish keyboard. When a letter was pressed the corresponding letter in the alphabet would blink, providing feedback to the subject. After this, the alphabet disappeared and a new trial would start. Each time the subject had carried out 100 new trials a message was displayed on the screen stating how many trials remained. On average, a trial took about 3 seconds. The stimulus conditions varying between trials were the identity of the letter and the exposure duration. All experimental blocks contained all stimulus conditions twice but the order of the stimulus conditions was randomly permuted for each block.

Subjects

Two Danish students of engineering as well as one of the authors served as subjects. The two students, subject MK and subject MH, were paid by the hour. Subject MK was a 24-year-old male with corrected to normal vision (contact lenses) and subject MH was a 21-year-old male, with corrected to normal vision (glasses). Subject AP was a 28-year old male with normal visual acuity. Subjects MK and MH were naïve about the purpose of the experiment, subject AP was not.

Stimulus display

The fixation point marker was a dot (\bullet), except for subject AP in Experiment 1, where the fixation point marker was a cross (+). The fixation point marker was shown at the centre of the screen, when the system was ready for a trial. The luminance of the point marker was: 8.9 Cd/m^2 . The background luminance of the screen was fixed at 45.7 Cd/m^2 in both Experiment 1 and 2.

The stimulus consisted of a single capital letter displayed at the fixation point. The letter could be any of the 26 letters of the English alphabet. The font used was New Courier. The letter was presented centrally (foveally) at the location of the fixation point. The letter subtended a visual angle of about 1.1° vertically and 1.1° horizontally.

Mask

The mask consisted of a binary image that was randomly generated for each trial. The procedure that was used for generating the instances of the mask was based on phase scrambling of the stimulus images in the Fourier domain. The procedure is described in detail in Appendix A. The mask was shown immediately after the stimulus was removed from the screen. The physical size of the mask was so that it would cover the area where all possible stimuli from A to Z had been presented, this means that the mask subtended about 1.3° vertically and 1.3° horizontally. The luminance of the binary mask was 45.7 Cd/m^2 in the light regions and 0 Cd/m^2 in the dark regions.

Report display

In the report display the entire alphabet from A to Z was displayed in a single row. The vertical angle between the vertical centre of the row of letters displayed and down to the bottom of the screen was 9° . The alphabet was printed in the same size as used for the stimuli letters. The entire

row of letters subtended about 1.1° vertically and 33° horizontally. The luminance of the letters in the report display was 32.2 Cd/m^2 .

Apparatus

The subject was seated in front of a computer-driven (NVIDIA GeForce 7950 GT) cathode ray screen (17" Flatron 915FT Plus) at a viewing distance of 57 cm in a darkened room. The viewing distance was chosen so that 1° of visual angle corresponded to approximately 1 cm on the screen. The refresh rate of the monitor was set to 200 Hz and the pixel resolution was 480 times 640. The monitor was pre-heated for at least half an hour before any experimental session was initiated. The experiment was written in Matlab™, using the Psychophysics Toolbox (Brainard, 1997, Pelli, 1997).

Analysis

We fitted the psychometric functions and TVA to our data by maximizing the likelihood using a quasi-Newtonian optimization routine provided by the Matlab™ optimization toolbox. We used a number of different starting points. We also added random noise to the parameters after convergence and then restarted the optimization. This was done to increase the chances of finding a global rather than a local minimum.

Experiment 1

In this experiment the negative Weber contrast of the stimulus letters was fixed at 0.0460. The experiment consisted of 65 sessions for each subject. Within a session each letter was presented two times at all exposure durations. There were 16 different exposure durations: 35, 40, 45, 50, 55, 60, 70, 80, 90, 100, 115, 130, 145, 165, 185 and 210 ms. With this setup a session contained a total of $2 \times 26 \times 16 = 832$ trials. The time to complete a session was about 40 minutes, after each session the

subject took a break for 10 minutes. A subject was not allowed to complete more than 5 sessions per day, and was told not to engage in any more sessions if they felt tired.

Experiment 2

In this experiment the stimulus was shown at 11 different contrast levels. Within each session the stimulus contrast level was fixed while it varied randomly between sessions. A total of 5 sessions contained the same contrast level, and within each session each condition was repeated two times. In total 55 sessions were completed. After each session the subject took a break for 10 minutes. The subject did not complete more than 6 sessions per day, and further did not engage in any more sessions if he felt tired.

For the stimuli, having the negative Weber contrast levels of 0.083, 0.046, 0.028, 0.020, and 0.018 the exposure durations were the same as the ones in Experiment 1, that is: 35, 40, 45, 50, 55, 60, 70, 80, 90, 100, 115, 130, 145, 165, 185 and 210 ms. The time to complete one of these sessions was about 40 minutes. This session type contained a total of $2 \times 26 \times 16 = 832$ trials. For the stimuli, having the negative Weber contrast levels of 0.370, 0.210 and 0.129, the exposure durations were: 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60 and 65 ms. The time to complete one of these sessions was about 30 minutes. This session type contained a total of $2 \times 26 \times 12 = 624$ trials. For the stimuli, having the negative Weber contrast levels of 1.000, 0.906 and 0.626 the exposure durations were: 0, 5, 10, 15, 20, 25, 30 and 35 ms. The time to complete one of these sessions was about 20 minutes. This session type contained a total of $2 \times 26 \times 8 = 416$ trials.

Results

In order to examine which psychometric function would describe the data when averaged across letter identity we fitted each of the six psychometric functions to the data from Experiment 1. The guessing rate was set to $\gamma = 1/26$ and the lapsing rate ($\lambda = 1-2\%$) was estimated from the data, as the average of the survivor function, $1-\psi$, (at 165, 185 and 210 ms). In Figure 1 we show the model error (the signed residual) as a function of exposure duration. Residuals are shown for the different psychometric functions and, comparing between the three different subjects, it is seen that the residuals vary systematically with exposure duration; notably, all of the psychometric functions overshoot around 40 to 60 ms and immediately after undershoot slightly less around 60-70 ms. It is seen from Figure 1 that the exponential is the least optimal model in the comparison, and further it is apparent that the residual for the squared-logistic and log-logistic are relatively small compared to the residuals of the other psychometric functions.

Figure 1

To further illustrate how well the psychometric functions fit the data, the proportion correct as a function of exposure duration averaged over all letter identities is shown in Figure 2. Also shown is the fit of the exponential psychometric function used in TVA and the fit of the squared-logistic psychometric function which we found provided the best fit (see Figure 1). For all subjects (AP, MH and MK) it is clear that for short exposure durations the proportion correct is slowly rising up until 20 % correct, which is not predicted by the exponential psychometric function, which rises abruptly in the beginning.

Figure 2

In order to study whether the hazard function of the psychometric function, F resembles the development of firing rates we show the hazard functions of the various models in Figure 3 with the empirical hazard rates. Across all subjects we consistently see that the empirical hazard rate rises smoothly and then falls to some lower level. Note, however, that the uncertainty of the hazard estimate increases as a function of exposure duration. The uncertainty bars (showing standard deviation) in the plot were obtained with the help of bootstrap analysis (Efron & Tibshirani, 1994), which consisted of fitting 1000 random re-samples of the data and calculating the standard deviation of these fits at each experimental condition. Note that since the hazard is calculated *between* neighbouring data points the exposure durations used in this plot are different than in Figure 1. Further for long exposure durations, as the psychometric function reaches ceiling, the hazard estimates become very imprecise or even infinite (Van Zandt, 2002), therefore we left these later estimates out of the plots.

Figure 3

To study to what extent the psychometric function depends on letter identity we also fitted each of the psychometric functions to the data from Experiment 1 *without* averaging over letter identities. The guessing rates were allowed to vary between letter identities, however the lapsing rates were not; i.e. these were the same as described above for the averaged data. To illustrate the parameter differences between letters, in Figure 4, we show parameter histograms as well as parameter scatter plots. These are shown for subject AP. Under the assumption that the psychometric function obtained by averaging over letter identities is the true psychometric function we applied bootstrapping (Efron & Tibshirani, 1994) to check if the variance in the parameters of the psychometric functions for the individual letter identities can be ascribed entirely to an effect of

random sampling or if it needs also to be ascribed to some systematic effect of letter identity. The bootstrapping consisted of fitting the psychometric function to 200 random re-samples of the data averaged over letter identities and then calculating the standard deviation for each of the model parameters based on the 200 model fits. The ovals in Figure 4 demark the bootstrap estimated confidence region to which we would expect 95 % of the letters, placed according to their individual parameters, to be located. Clearly many letters are located outside the ovals. This shows that the model parameters vary significantly with letter identity.

Figure 4

To examine how the psychometric function as a function of exposure duration depends on contrast, in Figure 5 we show proportion correct for the various contrast conditions in Experiment 2. For this experiment we note that the guessing rate, $\gamma = 1/26$, and further we shall assume that the lapsing rate, $\lambda = 0$. The fit by the squared-logistic psychometric function is also shown in Figure 5. The fit is very close for all the contrast levels that we investigated.

Figure 5

With the aim of investigating if any of the alternative psychometric functions can improve the ability of TVA to describe whole and partial report data, we inserted each of the six psychometric functions in TVA (See Appendix B). The six models were fitted to the data for each of the two subjects in (Shibuya & Bundesen, 1988). Clearly the squared-logistic fits closer than the exponential psychometric function; this is true for both subject MP (Figure 6.a) and subject HV (Figure 6.b). The squared-logistic, for example, is able to account for correct reports at very short

exposure durations because it does not rise abruptly in the beginning as does the exponential psychometric function used in (Shibuya & Bundesen, 1988).

Figure 6

Model feasibility for the different experiments is shown in Figure 7. The figures show AIC values (Akaike, 1974) and BIC values (Schwarz, 1978) for each psychometric function applied to each data set. AIC and BIC measures are shown summed over all subjects. Also included is the saturated model for reference. There are two graphs for Experiment 1, one for fitting the data averaged over letter identities (Figure 7.a) and one for fitting the data not averaging over the individual letter identities (Figure 7.b). In both graphs, comparing both AIC and BIC measures, we see that the exponential is the poorest and the squared-logistic and the log-logistic are the best psychometric functions. The model performance for the data in Experiment 2 is shown in Figure 7.c, we see that the ranking of the various models is quite similar to that found for Experiment 1, and further measures seem consistent between AIC and BIC measures. In Figure 7.d we show model performance with respect to modelling the whole and partial report data from (Shibuya & Bundesen, 1988). The ranking of the psychometric functions appear similar to what was seen for Experiments 1 and 2, and there appears to be a consistency between AIC and BIC measures. Finally Figure 7.e shows model performance with respect to modelling the data in (Bundesen & Harms, 1999). For this experiment, we again see that both the squared-logistic and the log-logistic are better models than the exponential psychometric function. However we see that the Weibull and gamma psychometric functions fit the data from this experiment better, which might appear curious if we compare this result with the results from modelling the other datasets. However, for most of the trials in (Bundesen & Harms, 1999) performance had reached ceiling level. Therefore, because of

fewer informative trials, we should be careful about putting too much weight on this dataset in the decision about which psychometric function is the most appropriate one to use.

An overview of model feasibility summed over all experiments, when *averaging* over letter identities, is shown in Figure 7.f. The graph shows the sum of AIC and BIC measures for the data in Experiment 1 and 2 as well as the data in (Shibuya & Bundesen, 1988) and (Bunesen & Harms, 1999). To ensure consistency with the way the other datasets were modelled, for Experiment 1 we include the AIC and BIC values obtained after fitting the data, when data was averaged over letter identities. Thereby the AIC and BIC values found when modelling the individual letters in Experiment 1 (cf. Figure 7.b) are not included in the sums shown in Figure 7.f. From Figure 7.f we see that the three-parameter squared-logistic is the best model with the two-parameter log-logistic model as the runner-up. The two-parameter exponential distribution is clearly the poorest model. The ranking appears consistent over AIC and BIC measures, and also for both AIC and BIC, we see that the exponential is poorer, while the squared-logistic is better, than the saturated model. For the most optimal psychometric function for *individual* letters we refer to Figure 7.b, which shows that although the AIC and BIC measures disagree whether the squared-logistic or the log-logistic psychometric function provides the best fit the two measures do agree that the squared-logistic and log-logistic provide a much better fit than the exponential psychometric function.

Figure 7

Discussion

TVA (Bundesen, 1990) has been successfully used to model data from whole and partial report type of experiments (Shibuya & Bundesen, 1988). For this, TVA used the exponential distribution as the underlying psychometric function; however since many other psychometric functions exist, we wondered whether a more optimal psychometric function could be found. In order to answer that question we conducted Experiment 1, which was a single-letter identification experiment, similar to that of (Bundesen & Harms, 1999), in which exposure duration was varied at a single fixed contrast level. Our first idea was to generalise the psychometric function by simply including an additional, third parameter. The Weibull, the gamma and the ex-Gaussian all represented simple generalisations of the exponential psychometric function, and all of these proved to be better models as measured by AIC and BIC comparisons.

From NTVA (Bundesen et al., 2005) comes the prediction that the hazard rate of the psychometric function should follow the firing rates of neurons in the visual cortex, which have been shown to develop non-monotonically over time (Albrecht et al., 2002). Therefore we included two additional psychometric functions which both have a non-monotonic hazard function: the log-logistic which is a well-known distribution and the squared-logistic, which is a biologically inspired distribution we developed ourselves. The three-parameter squared-logistic generally produced the best results, but the two-parameter log-logistic came out as the runner-up, comparing all six different psychometric functions (see Figure 7).

To illustrate how well the different psychometric functions fit the single-letter identification data from Experiment 1 in Figure 1.a-c we showed the model error as a function of exposure duration for the three different subjects. Comparing across the three subjects, it is clear that the error function

develops systematically over time. Even for the best fitting psychometric function, that is the squared-logistic, it appears that first there is an overshoot around 40-60 ms and then an undershoot around 60-70 ms. In Figure 3 we compared the empirically estimated hazard functions with the hazard functions of the various psychometric functions used. When seen across all subjects in Experiment 1, it is seen (from both Figure 1 and Figure 3) that we find that there is a small, but consistent, systematic misfit as a function of exposure duration, and so we invite future studies to find an even better suited model than we did for characterising the temporal development of identification accuracy.

To the best of our knowledge, this study is unique in the sense that it contains enough trials per subject to allow us to accurately fit the psychometric function for each individual letter. For instance in the earlier single-letter identification study that we mentioned (Bundesen & Harms, 1999) averages over all stimuli letters were used. Our study offers novel knowledge about how dependent the parameters of the psychometric function are on letters identity. Making use of bootstrapping (Efron & Tibshirani, 1994) we found that the parameters of the estimated psychometric functions of the individual letters vary more than would be expected from random sampling variance alone. It is however common to average the psychometric function over letter identities, and therefore it is relevant to ask whether it is reasonable to use the same type of psychometric function, both when fitting individual letters, as well as when fitting the data averaged over letters identities. Our results showed that the log-logistic and squared-logistic psychometric functions are optimal in both of these two situations. This means that although the model parameters depend on letter identity, the type of psychometric function does not. Therefore, it still seems reasonable in many types of experiments to average over letter identities to reduce the demand on the total number of trials.

Worried that our search for an optimal psychometric function would suffer from being contrast-specific, we conducted Experiment 2 to verify the performance of the various psychometric functions at a number of different stimulus contrast levels. The result of fitting the data from Experiment 2 is illustrated in Figure 7.c. For Experiment 2 the ranking of the psychometric functions was quite similar to the one we had previously seen for Experiment 1. In Figure 5 we showed how well the squared-logistic psychometric function fits the data from Experiment 2.

Interestingly the two best-fitting models we found, namely the log-logistic and the squared-logistic both have a non-monotonic hazard function in line with the prediction of Bundesen and Habekost (2008). Other distributions including the Cauchy and the log-normal also have a non-monotonic hazard function, but these distributions did not provide as good fits (not shown) as the log-logistic and squared-logistic did. The log-logistic distribution is a special case of the Burr distribution, which again is a special case of the generalized beta distributions of the second kind (Bookstaber & McDonald, 1987). Little improvement in fitting was found when testing these generalizations (not shown), which is why here we do not go into more detail on these functions. Non-monotonic hazard functions are described in (Aalen & Gjessing, 2001).

After insertion of each of the six psychometric functions into TVA (Bunesen, 1990), further described in Appendix B, the three-parameter squared-logistic enabled the closest fits (see Figure 7.d) to the data from (Shibuya & Bundesen, 1988). A close runner-up was the log-logistic psychometric function.

In general over all datasets that we fitted (single-letter as well as whole and partial report) it is clear (see Figure 7) that the squared-logistic is the most optimal psychometric function of the ones that we have considered. An alternative to using the three-parameter squared-logistic psychometric function is to use the log-logistic psychometric function, which comes in as a close runner-up. Despite having only two parameters, the log-logistic still provides very good fits to the data. It is worth noticing that Bundesen's exponential psychometric function also had two parameters; however the log-logistic provides much better fits.

We investigated visual letter identification as a function of exposure duration and describe the squared-logistic, a psychometric function we found no previous accounts for, and which we developed motivated by NTVA (Bundesen et al., 2005) and single-neuron studies by (Albrecht et al., 2002). Both the squared-logistic and the well-known log-logistic can model a non-monotonic hazard function and both of these two psychometric functions fit well to experimental data from single-letter identification experiments; finally inserted into TVA (Bundesen, 1990), both psychometric functions improve fits to data from whole and partial report type of experiments. For all datasets that we modelled we found that the three-parameter squared-logistic and the two-parameter log-logistic were clearly better models than the two-parameter exponential psychometric function, which has until now been used with TVA.

Acknowledgements

The authors thank Claus Bundesen, Lars Kai Hansen and Søren Kyllingsbæk for helpful comments on various issues concerning the subject of this article. We also wish to thank Hitomi Shibuya, Lisbeth Harms and Claus Bundesen for access to the data from (Shibuya & Bundesen, 1988) and (Bundenen & Harms, 1999). This research was funded by The Oticon Foundation (AP) and The Danish Council for Strategic Research (TSA).

References

- Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transactions On Automatic Control*, *AC-19*(6), 716–723.
- Albrecht, D. G., Geisler, W. S., Frazor, R. A., & Crane, A. M. (2002). Visual cortex neurons of monkeys and cats: Temporal dynamics of the contrast response function. *Journal Of Neurophysiology*, *88*(2), 888–913.
- Alexander, K. R., Xie, W., & Derlacki, D. J. (1997). Visual acuity and contrast sensitivity for individual Sloan letters. *Vision Research*, *37*(6), 813–819.
- Baldock, M. R. J., Mathias, J., McLean, J., & Berndt, A. (2007). Visual attention as a predictor of on-road driving performance of older drivers. *Australian Journal Of Psychology*, *59*(3), 159–168.
- Behrmann, M., Nelson, J., & Sekuler, E. B. (1998). Visual complexity in letter-by-letter reading: "Pure" alexia is not pure. *Neuropsychologia*, *36*(11), 1115–1132.
- Bookstaber, R. M., & McDonald, J. B. (1987). A General Distribution For Describing Security Price Returns. *Journal Of Business*, *60*(3), 401–424.

- Brüederl, J., & Diekmann, A. (1995). The Log-Logistic Rate Model - 2 Generalizations With An Application To Demographic-Data. *Sociological Methods & Research*, 24(2), 158–186.
- Bundesen, C. (1990). A Theory of Visual Attention. *Psychological Review*, 97(4), 523–547.
- Bundesen, C. (1993). The Relationship Between Independent Race Models And Luce Choice Axiom. *Journal Of Mathematical Psychology*, 37(3), 446–471.
- Bundesen, C. (1998). A computational theory of visual attention. *Philosophical Transactions Of The Royal Society Of London Series B-Biological Sciences*, 353(1373), 1271–1281.
- Bundesen, C., & Habekost, T. (2008). *Principles of Visual Attention: Linking Mind and Brain* (1st ed.). Oxford University Press, USA.
- Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2005). A Neural Theory of Visual Attention: Bridging Cognition and Neurophysiology. *Psychological Review*, 112(2), 291–328.
- Bundesen, C., & Harms, L. (1999). Single-letter recognition as a function of exposure duration. *Psychological Research-Psychologische Forschung*, 62(4), 275–279.
- Chen, C. H., & Chien, Y. H. (2005). Reading Chinese text on a small screen with RSVP. *Displays*, 26(3), 103–108.
- Cheong, A. M. Y., Legge, G. E., Lawrence, M. G., Cheung, S. H., & Ruff, M. A. (2007). Relationship between slow visual processing and reading speed in people with macular degeneration. *Vision Research*, 47(23), 2943–2955.
- Chien, Y. H., Chen, C. H., & Wei, W. L. (2008). Effects of dynamic display, presentation method, speed, and task type on reading comprehension of wristwatch screens. *Displays*, 29(5), 471–477.
- Di Lollo, V., von Mühlelen, A., Enns, J. T., & Bridgeman, B. (2004). Decoupling stimulus duration from brightness in metacontrast masking: Data and models. *Journal Of Experimental Psychology-Human Perception And Performance*, 30(4), 733–745.

- Diekmann, A. (1992). The Log-Logistic Distribution As A Model For Social Diffusion-Processes. *Journal Of Scientific & Industrial Research*, 51(3), 285–290.
- Efron, B., & Tibshirani, R. (1994). *An Introduction to the Bootstrap* (1st ed.). Chapman & Hall/CRC.
- Habekost, T., & Starrfelt, R. (2008). Visual attention capacity: A review of TVA-based patient studies. *Scandinavian Journal of Psychology*, 50(1), 23–32. doi: 10.1111/j.1467-9450.2008.00681.x.
- Lansky, P., Pokora, O., & Rospars, J. P. (2007). Stimulus-response curves in sensory neurons: how to find the stimulus measurable with the highest precision. *Advances in Brain, Vision, and Artificial Intelligence. Proceedings Second International Symposium, BVAI 2007*, 338-349.
- Luce, R. D. (1991). *Response Times: Their Role in Inferring Elementary Mental Organization*. Oxford University Press, USA.
- Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of Optical Society America A*, 2(9), 1508–1532.
- Pelli, D. G. (1987). On the relation between summation and facilitation. *Vision Research*, 27(1), 119–123.
- Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research*, 46(28), 4646–4674. doi: 10.1016/j.visres.2006.04.023.
- Pelli, D. G., & Tillman, K. A. (2007). Parts, Wholes, and Context in Reading: A Triple Dissociation. *PLoS ONE*, 2(8), e680. doi: 10.1371/journal.pone.0000680.
- Rasinski, T. V. (2000). Speed does matter in reading. *Reading Teacher*, 54(2), 146–151.
- Richardson, E. D., & Marottoli, R. A. (2003). Visual attention and driving behaviors among community-living older persons. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 58(9), M832–M836.

- Schwarz, G. (1978). Estimating Dimension Of A Model. *Annals Of Statistics*, 6(2), 461–464.
- Shibuya, H. (1994, September 10). Time course of elementary processes underlying visual selective attention: Shape of hazard functions. Lisbon, Portugal.
- Shibuya, H., & Bundesen, C. (1988). Visual Selection from Multielement Displays: Measuring and Modeling Effects of Exposure Duration. *Journal of Experimental Psychology Human Perception Performance*, 14(4), 591–600.
- Strasburger, H. (2001). Invariance of the psychometric function for character recognition across the visual field. *Perception & Psychophysics*, 63(8), 1356–1376.
- Treutwein, B., & Strasburger, H. (1999). Fitting the psychometric function. *Perception & Psychophysics*, 61(1), 87–106.
- Van Breukelen, G. J. P. (1995). Psychometric And Information-Processing Properties Of Selected Response-Time Models. *Psychometrika*, 60(1), 95–113.
- Van Zandt, T. (2002). Steven's Handbook of Experimental Psychology. In S. S. Stevens & H. E. Pashler (Eds.), (3rd ed.). John Wiley and Sons.
- Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception & Psychophysics*, 63(8), 1314–1329.
- Aalen, O. O., & Gjessing, H. K. (2001). Understanding the shape of the hazard rate: A process point of view. *Statistical Science*, 16(1), 1–14.

Appendix A: Generating a random Fourier mask

Here we explain the method we used for generating a new random mask M for each trial. The idea behind the method is to generate a binary mask that should have a power spectrum that is close to the average power spectrum of the stimulus images used. Further for each trial the phase content of the mask should be randomised.

Our starting point is that we have N binary stimulus images of size X_1 times X_2 , the n 'th image $f_n(x_1, x_2)$ has pixel coordinates x_1 and x_2 . Noting that w_1 and w_2 are the pixel coordinates in frequency space, the discrete Fourier transform of the n 'th stimulus image can be written as:

$$F_n(w_1, w_2) = \sum_{x_2=1}^{X_2} \sum_{x_1=1}^{X_1} f_n(x_1, x_2) e^{-i \cdot w_1 \cdot x_1} e^{-i \cdot w_2 \cdot x_2}$$

The average Fourier transform F of the N stimulus images can be calculated as:

$$\hat{F}(w_1, w_2) = \frac{1}{N} \sum_{n=1}^N F_n(w_1, w_2)$$

Let Θ be a matrix that has size X_1 times X_2 . We now define the randomly phase-shifted average Fourier transform as:

$$F^\phi(w_1, w_2) = |\widehat{F}(w_1, w_2)| \begin{cases} 1 & \text{if } w_1 = 0 \wedge w_2 = 0 & (DC) \\ 1 & \text{elseif } w_1 = \frac{1}{2X_1} \wedge w_2 = \frac{1}{2X_2} & \left(\frac{f_s}{2}\right) \\ \exp(\Theta) & \text{else} \end{cases}$$

We note that entries in Θ are chosen randomly for each trial and subject to the constraint that F^ϕ should be a Hermitian matrix. The entries for Θ are drawn in conjugated pairs with zero real and random imaginary part. Also note that for the later equation the DC-frequency and the half sampling frequency $f_s/2$ were not phase-scrambled, because both frequency components should be kept real.

The discrete inverse Fourier transform of F^ϕ can be written as:

$$f^\phi(x_1, x_2) = \sum_{w_1=-\pi}^{\pi} \sum_{w_2=-\pi}^{\pi} F^\phi(w_1, w_2) e^{i \cdot w_1 \cdot x_1} e^{i \cdot w_2 \cdot x_2}$$

We now define the rounded average number of pixels ν in the N stimulus images that takes on the value 1 rather than 0 as

$$\nu = \text{round} \left(\frac{1}{N} \sum_{n=1}^N \sum_{x_1=1}^{X_1} \sum_{x_2=1}^{X_2} f_n(x_1, x_2) \right)$$

We now perform a threshold operation so that the v elements in f_ϕ that have the largest values are set to one in M and the rest of the elements are set to zero; i.e. if we define as S_v the set of elements in f_ϕ that have the v largest values, then the random Fourier mask M is found as:

$$M(x_1, x_2) = \begin{cases} 1 & \text{if } f^\phi(x_1, x_2) \in S_v \\ 0 & \text{else} \end{cases}$$

Appendix B: The theory of visual attention and the generalized fixed-capacity independent race model

TVA (Bundesen, 1990) provides an account of performance in whole and partial report. In TVA any target or distractor is denoted as an *element*, and further it is assumed that a subject only correctly identifies the elements that are stored in visual short-term memory. Further, subjects only obtain a chance to *store* an element if the element is *encoded*.

Encoding

The hazard rate that a particular element, i , is encoded into VSTM is proportional to how large a portion of processing resources the element receives. Any element in the visual field S receives a certain portion v_i of the total processing capacity C , which is assumed to be invariant with respect to the number of elements in the display:

$$C = \sum_{i \in S} v_i$$

It serves as a simplification to assume homogeneity of the visual display (Shibuya & Bundesen, 1988). This appears a reasonable assumption as long as all elements have the same size, the same contrast, the same eccentricity etc. The homogeneity assumption means that all targets receive the same amount of processing resources denoted v_t . In the same way all distractors receive the same amount v_d of processing resources which is proportionally smaller than the amount that targets receive. The ratio of processing resources α is defined as:

$$\alpha = \frac{v_d}{v_t}$$

Let us assume a homogenous display that contains T targets and D distractors. The processing resources v_t of any target is then given by:

$$C = \sum_{i \in S} v_i \Rightarrow C = Tv_t + Dv_d = Tv_t + \alpha Dv_t = (T + \alpha D)v_t \Leftrightarrow v_t = \frac{C}{T + \alpha D}$$

Storage

As we have already seen, TVA describes the factors that determine the probability that any given target in a multi-element visual display is encoded into visual short-term memory, however as VSTM typically has only about 3-4 storage places, TVA assumes that not all elements that become encoded are actually stored. Bundesen (1990) assumes that the occupation of places occurs through a so-called race, that is, any newly encoded element will lead to immediate occupation of one storage place if and only if there is still any storage place left in the VSTM.

Let us define that f and F and are the probability density and the distribution function for target encoding respectively. Similarly we also define that g and G are the probability density and the distribution function for distractor encoding.

According to the fixed-capacity independent race model (FIRM, Shibuya & Bundesen, 1988) the probability of a score of j (targets reported) from a display containing T targets and D distractors exposed for τ seconds can be written as:

$$P(j; T, D, \tau) = P_1 + P_2 + P_3$$

where P_1 is the probability that the score equals j and the total number of elements (targets and distractors) entering VSTM is *less* than K . The number of distractors entering VSTM is denoted by m and $m \leq \min(D, K-j-1)$. If $j=K$, $P_1=0$; otherwise:

$$P_1 = \binom{T}{j} [F(\tau)]^j [1-F(\tau)]^{T-j} \cdot \sum_{m=0}^{\min(D, K-j-1)} \binom{D}{m} [G(\tau)]^m [1-G(\tau)]^{D-m}$$

and P_2 is the probability that the score equals j and the total number of elements *equals* K and the K^{th} element entering the VSTM is a *target*. The number of distractors entering VSTM denoted by m is always $K-j$. If $j=0$, or $j < K-D$, $P_2=0$; otherwise:

$$P_2 = \int_0^{\tau} \binom{T-1}{j-1} [F(t)]^{j-1} [1-F(t)]^{T-j} \binom{D}{m} [G(t)]^m [1-G(t)]^{D-m} \binom{T}{1} f(t) dt$$

and P_3 is the probability that the score equals j and the total number of elements *equals* K and the K^{th} element entering the VSTM is a *distractor*. The number of distractors entering VSTM denoted by m is always $K-j$. If $j=K$, or $j < K-D$, $P_3=0$; otherwise:

$$P_3 = \int_0^{\tau} \binom{T}{j} [F(t)]^j [1-F(t)]^{T-j} \binom{D-1}{m-1} [G(t)]^{m-1} [1-G(t)]^{D-m} \binom{D}{1} g(t) dt$$

Shibuya and Bundesen (1988) derived explicit score probabilities under the assumption that encoding proceeds as a homogenous Poisson process. This corresponds to assuming (ignoring their suggested temporal offset) that the hazard rates are constant over time. Our contribution is to allow the hazard rates to be time-varying, although we still assume that the hazard rates (for the different elements presented) are mutually proportional functions of time (cf. Bundesen, 1990, 1993, 1998). The way we relax this assumption is explained in the following where we derive a set of generalised FIRM equations. For the generalized FIRM equations, which are characterized by a non-homogenous Poisson process for encoding, with hazard $\lambda_i(t)$ and cumulative hazard function $\Lambda_i(t)$, we can write the following probability distribution function for target encoding:

$$F(t) = 1 - \exp(-\Lambda_i(t))$$

By differentiating the probability distribution using the chain rule we arrive at the probability density function for target encoding:

$$f(t) = \exp(-\Lambda_i(t)) \lambda_i(t)$$

Similarly, for distractors we can formulate the probability density function g and the probability distribution G in terms of the hazard $\lambda_d(t)$ and the cumulative hazard function $\Lambda_d(t)$ for distractor encoding.

We are now able to derive a set of explicit score probabilities that are valid under the generalized FIRM conditions. By inserting the expressions for $F(t)$ and $G(t)$ it is straightforward to calculate the probability P_I when the cumulative hazard functions $\Lambda_i(t)$ and $\Lambda_d(t)$ are known:

$$P_1 = \binom{T}{j} [1 - \exp(-\Lambda_t(\tau))]^j [\exp(-\Lambda_t(\tau))]^{T-j} \sum_{m=0}^{\min(D, m-1)} \binom{D}{m} [1 - \exp(-\Lambda_d(\tau))]^m [\exp(-\Lambda_d(\tau))]^{D-m}$$

Deriving the expression for P_2 is a little more complex:

$$\begin{aligned} P_2 &= \int_0^\tau \binom{T-1}{j-1} [F(t)]^{j-1} [1-F(t)]^{T-j} \binom{D}{m} [G(t)]^m [1-G(t)]^{D-m} \binom{T}{1} f(t) dt \\ &= \binom{T-1}{j-1} \binom{D}{m} \binom{T}{1} \int_0^\tau [1 - \exp(-\Lambda_t(t))]^{j-1} \exp(-\Lambda_t(t))^{T-j} [1 - \exp(-\Lambda_d(t))]^m \exp(-\Lambda_d(t))^{D-m} \exp(-\Lambda_t(t)) \lambda_t(t) dt \\ &= \binom{T-1}{j-1} \binom{D}{m} \binom{T}{1} \int_0^\tau \exp(-\Lambda_t(t) \{T-j+1+\alpha(D-m)\}) [1 - \exp(-\Lambda_t(t))]^{j-1} [1 - \exp(-\Lambda_d(t))]^m \lambda_t(t) dt \\ &= \binom{T-1}{j-1} \binom{D}{m} \binom{T}{1} \int_0^\tau \exp(-\Lambda_t(t) \{T-j+1+\alpha(D-m)\}) \sum_{a=0}^{j-1} \binom{j-1}{a} \{-\exp(-\Lambda_t(t))\}^a \sum_{b=0}^m \binom{m}{b} \{-\exp(-\Lambda_d(t))\}^b \lambda_t(t) dt \\ &= \binom{T-1}{j-1} \binom{D}{m} \binom{T}{1} \sum_{a=0}^{j-1} \binom{j-1}{a} \left\{ \sum_{b=0}^m \binom{m}{b} (-1)^{a+b} \int_0^\tau [\exp(-\Lambda_t(t)) \{(T-j+1+a)+\alpha(D-m+b)\}] \lambda_t(t) dt \right\} \\ &= \binom{T-1}{j-1} \binom{D}{m} \binom{T}{1} \sum_{a=0}^{j-1} \binom{j-1}{a} \left\{ \sum_{b=0}^m \binom{m}{b} (-1)^{a+b} \frac{\exp(-\Lambda_t(\tau)) \{(T-j+1+a)+\alpha(D-m+b)\} - 1}{-(T-j+1+a)-\alpha(D-m+b)} \right\} \end{aligned}$$

In the first step, we insert the expressions for $F(t)$, $G(t)$ and $f(t)$ while the second step is a simple reduction. The third step uses the binomial expansion while the fourth step is again a simple reduction. The fifth step uses integration by substitution, noticing that $\partial A / \partial t = \lambda$, to arrive at the final expression. Similarly, we can derive the following expression for P_3 :

$$P_3 = \binom{T}{j} \binom{D-1}{m-1} \binom{D}{1} \sum_{a=0}^j \left\{ \binom{j}{a} \sum_{b=0}^{m-1} \left[\binom{m-1}{b} (-1)^{a+b} \alpha \frac{\exp(-\Lambda_t(\tau) \{(T-j+a) + \alpha \cdot (D-K+j+1+b)\}) - 1}{\{(T-j+a) - \alpha \cdot (D-K+j+1+b)\}} \right] \right\}$$

From these three expressions we can calculate the score probability, $P(j; T, D, \tau) = P_1 + P_2 + P_3$, from the cumulative hazard function, $\Lambda_t(t)$. To find the cumulative hazard function, $\Lambda_t(t)$, from a distribution function, $F(t)$, we note that it can be calculated as the negative logarithm of the survivor function (Luce, 1991), i.e.:

$$\Lambda_t(t) = -\log(1 - F(t))$$

Thus, all distribution functions including all psychometric functions known to us can be inserted into TVA using the above derivations. Note that in the case of a single target letter, the distribution function $F(t)$ is the psychometric function.

Finally, let us note that our assumption of an in-homogenous Poisson process for visual encoding; rather than a homogenous one as was assumed in (Bundesen, 1990); does not necessarily conflict with the idea of a total processing capacity C , if one assumes that this is no longer constant but rather varying in time. In this way we can use the same formulas for dividing processing resources between elements as used in (Bundesen, 1990). Also we note that the processing rate, ν , which was previously a constant, is now a time-varying function $\lambda(t)$ that corresponds to the hazard of encoding an element.

Figure legends

Figure 1

Residuals, plotted as a function of exposure duration, from Experiment 1. Error bars – too small to be distinguished clearly – show the standard error of the mean. There is one graph for each subject: a) AP, b) MH and c) MK.

Figure 2

Proportion correct in Experiment 1 averaged over letter identities. Error bars show the standard error of the mean. The fit of the exponential and the squared-logistic psychometric functions are shown as well. There is one graph for each subject: a) AP, b) MH and c) MK.

Figure 3

The hazard function plotted against exposure duration. Grey squares represent the empirical hazard function, which was estimated directly from the data. Error bars indicate the standard deviation as estimated by a bootstrap procedure. The hazard functions of the various psychometric functions are displayed as coloured lines. There is one figure for each subject: a) AP, b) MH and c) MK.

Figure 4

Scatter plots and histograms for the parameters of the log-logistic model fitted to the data from subject AP in Experiment 1. In the off-diagonal windows we see parameter pairs plotted against each other, while in the diagonal windows we see the parameter histograms. Each letter from A to Z was fitted individually giving each letter its own set of model parameters.

Figure 5

Proportion correct averaged over letter identities (as circles) for subject AP in Experiment 2. Error bars show standard deviation of the mean. Also shown (as lines) are the fits of the squared-logistic psychometric function. The legend shows the negative Weber contrast that was used. The horizontal dashed line shows the guessing level.

Figure 6

Cumulative score distributions from a whole and partial report experiment (Shibuya & Bundesen, 1988). Circular markers correspond to the proportion of scores of j or more (correctly reported targets) showed as a function of exposure duration. The legend for the scores is: blue: $j=1$, green: $j=2$, red: $j=3$, turquoise: $j=4$ and violet $j=5$. Each graph shows data for a certain combination of the number of targets, T , and the number of distractors, D . The dotted lines represent the fit of the exponential psychometric function inserted into TVA. Solid lines represent the fit of the squared-logistic psychometric function inserted into TVA. There are two sub-figures, one for each subject: a) MP and b) HV.

Figure 7

Model feasibility cumulated over all subjects. In the different figures we see how successful the different psychometric functions are at modelling various datasets. The measures shown are AIC values (Akaike, 1974) and BIC values (Schwarz, 1978). In a) we see the results from fitting the functions to the data from Experiment 1 averaged over letter identities. In b) we see the results from fitting the functions to the data from Experiment 1 without averaging over letter identities. In c) we see the result from fitting the functions to the data from Experiment 2. In d) we see the result from fitting our models to the data from the whole and partial report experiment in (Shibuya & Bundesen, 1988). In e) we see the results from fitting our functions to the single-letter identification data from (Bundesen & Harms, 1999). Finally, in f) we see AIC and BIC measures cumulated over all datasets from the various experiments, including only the fit to the average single-letter data (i.e. not the fit to the individual letters) for Experiment 1.

Figures

Figure 1

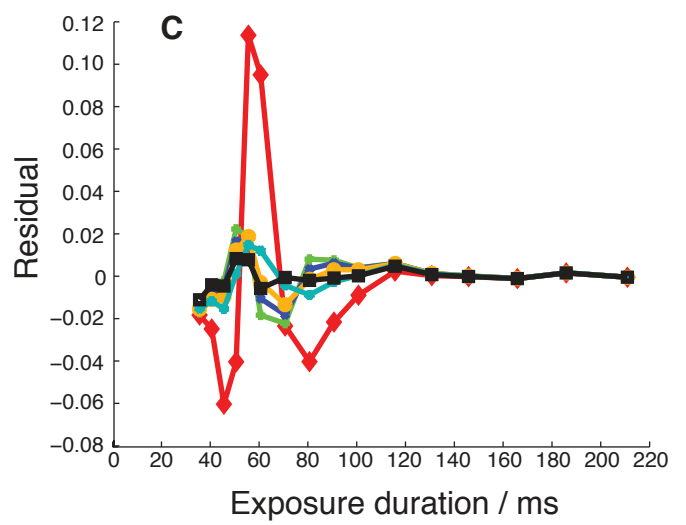
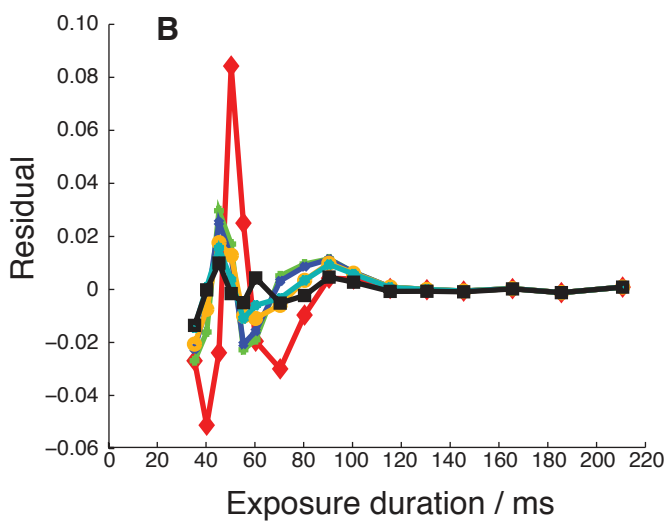
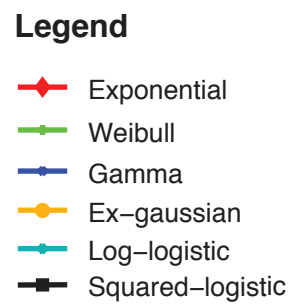
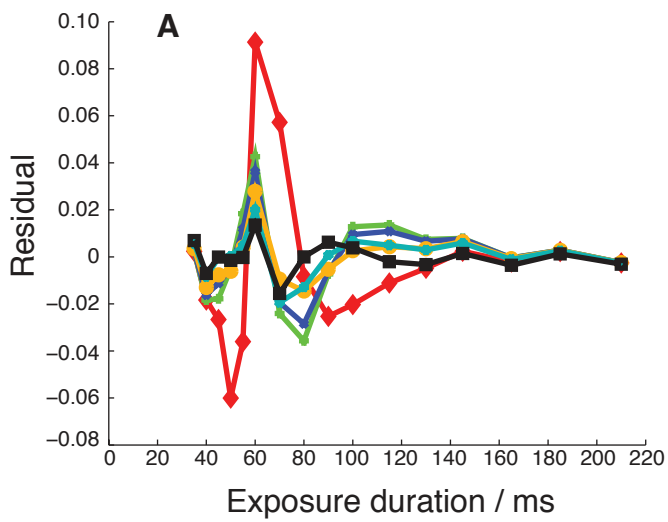
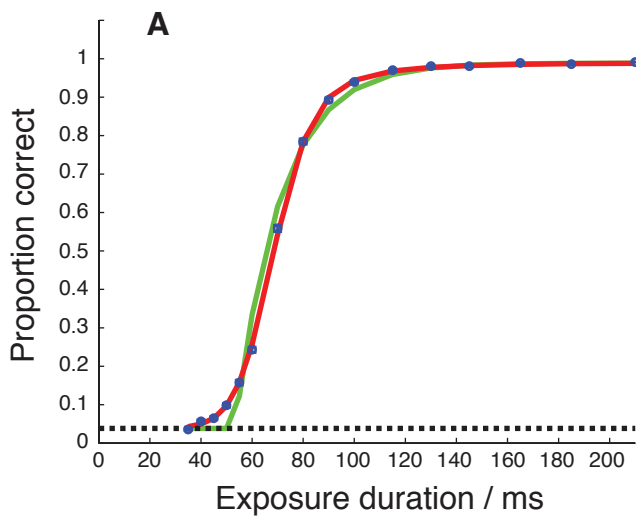


Figure 2



Legend

- Averaged response proportions
- Exponential
- Squared-logistic
- ⋯ Guessing rate

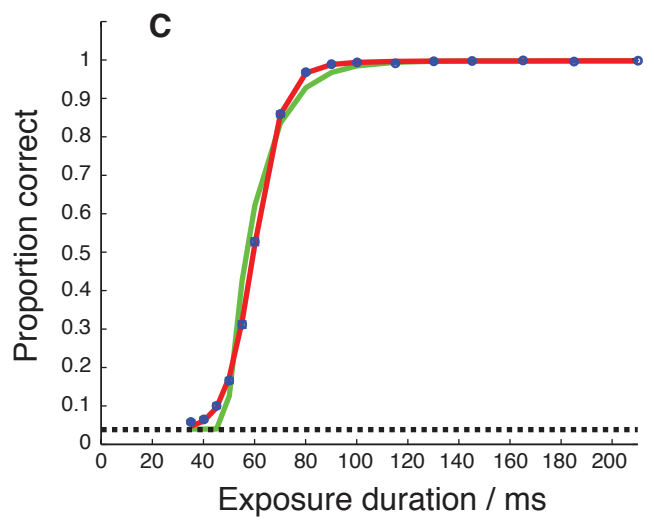
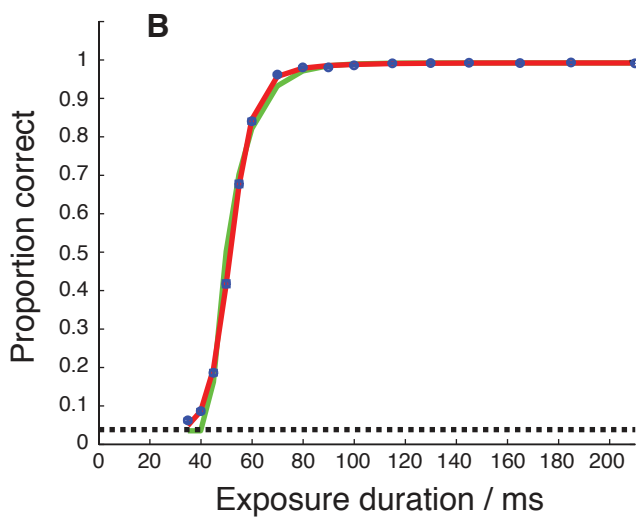
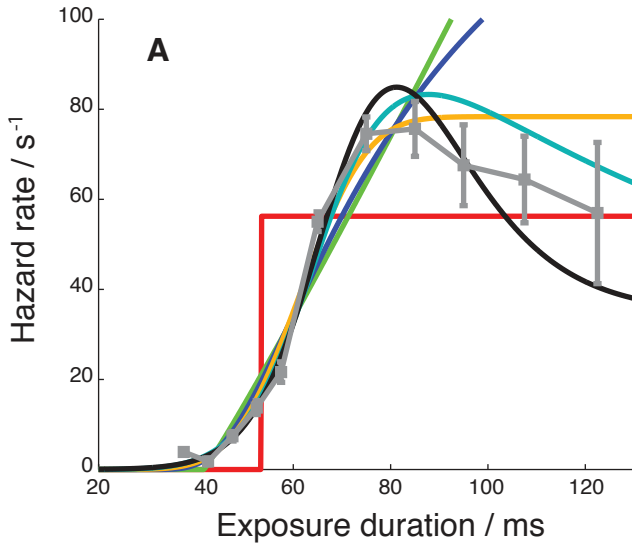


Figure 3



Legend

- Exponential
- Weibull
- Gamma
- Ex-gaussian
- Log-logistic
- Squared-logistic
- Empirical hazard rate

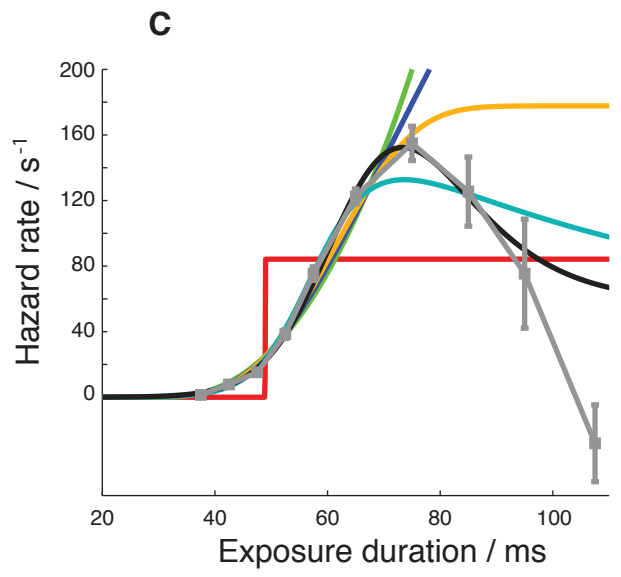
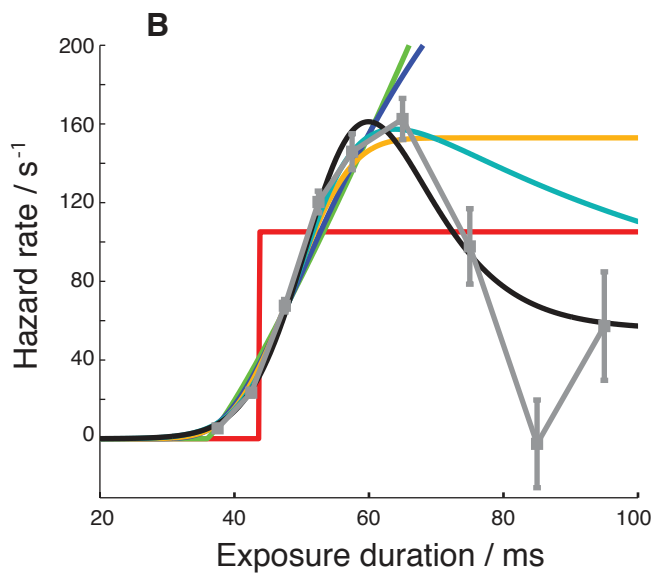


Figure 4

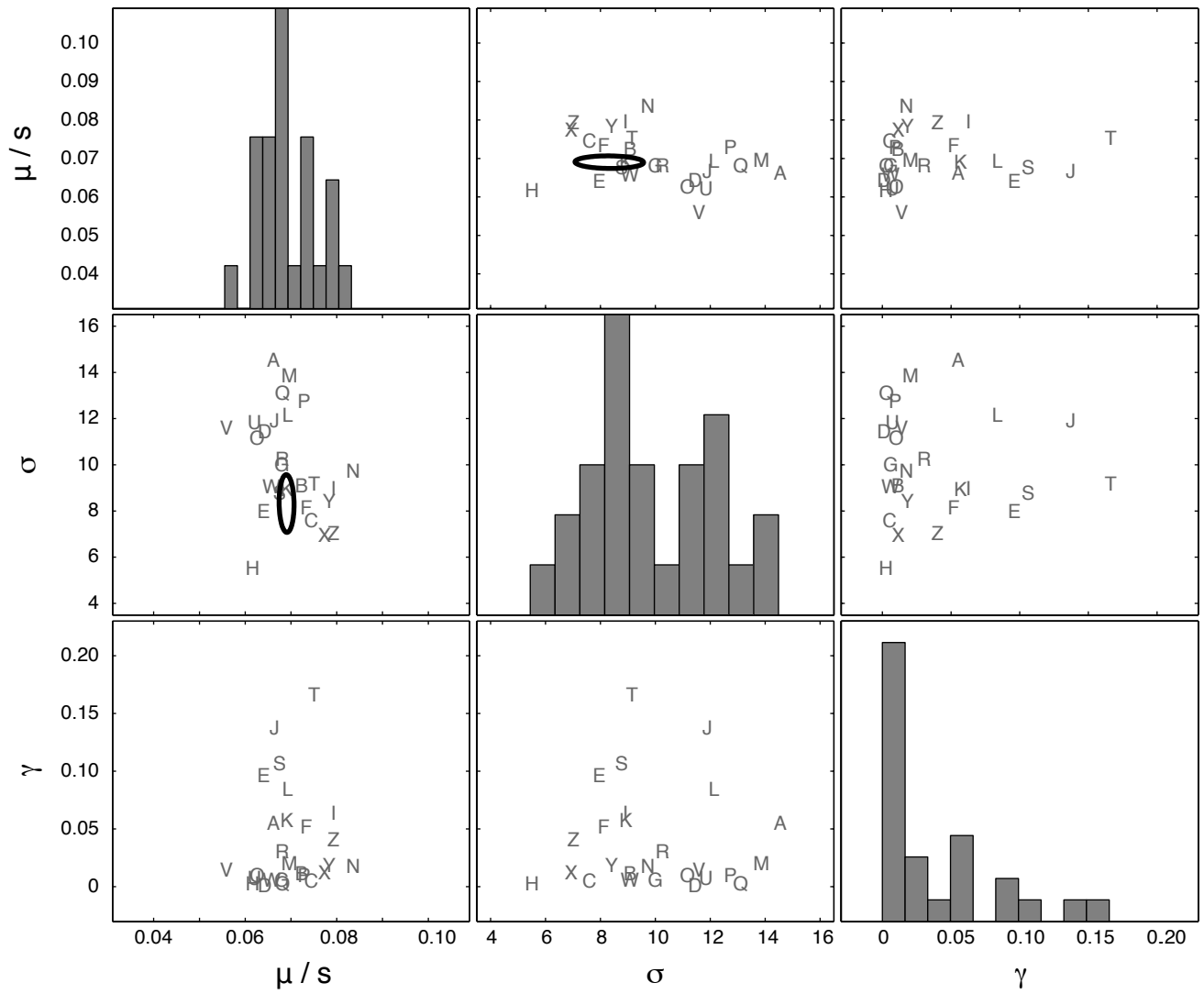


Figure 5

Single object report as a function of exposure duration

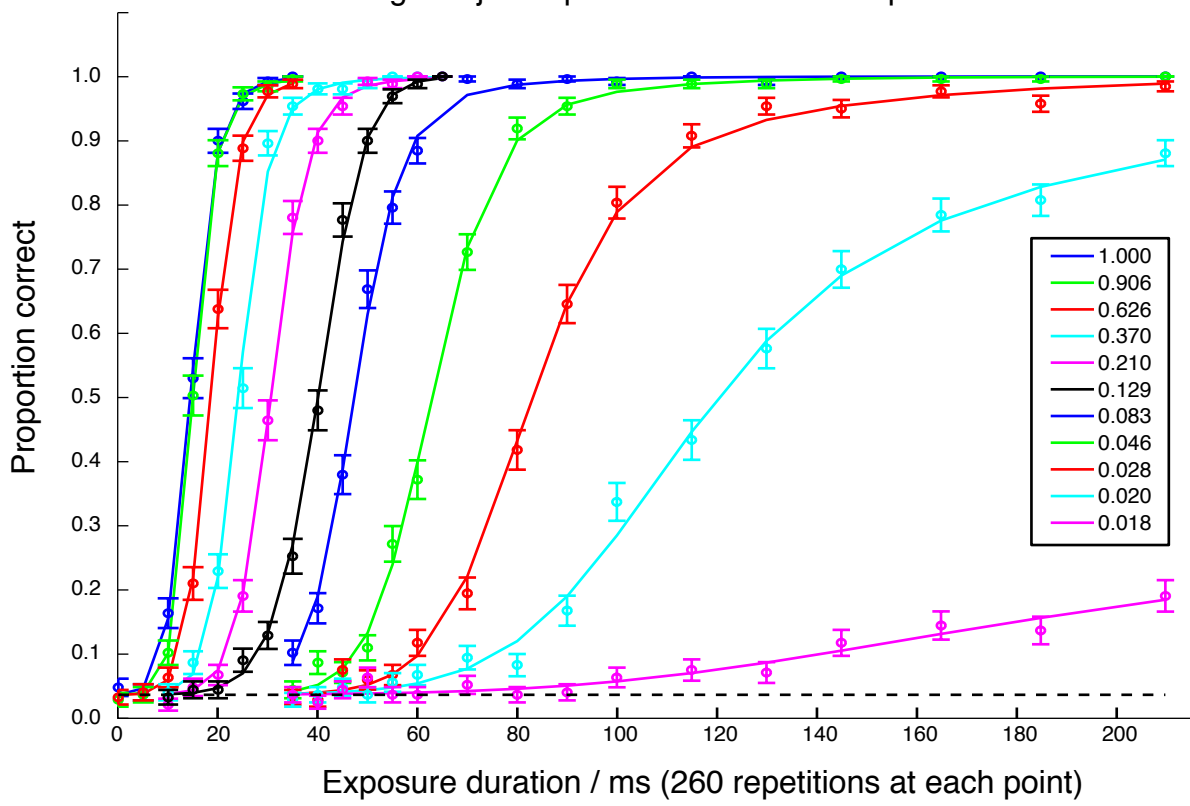


Figure 6

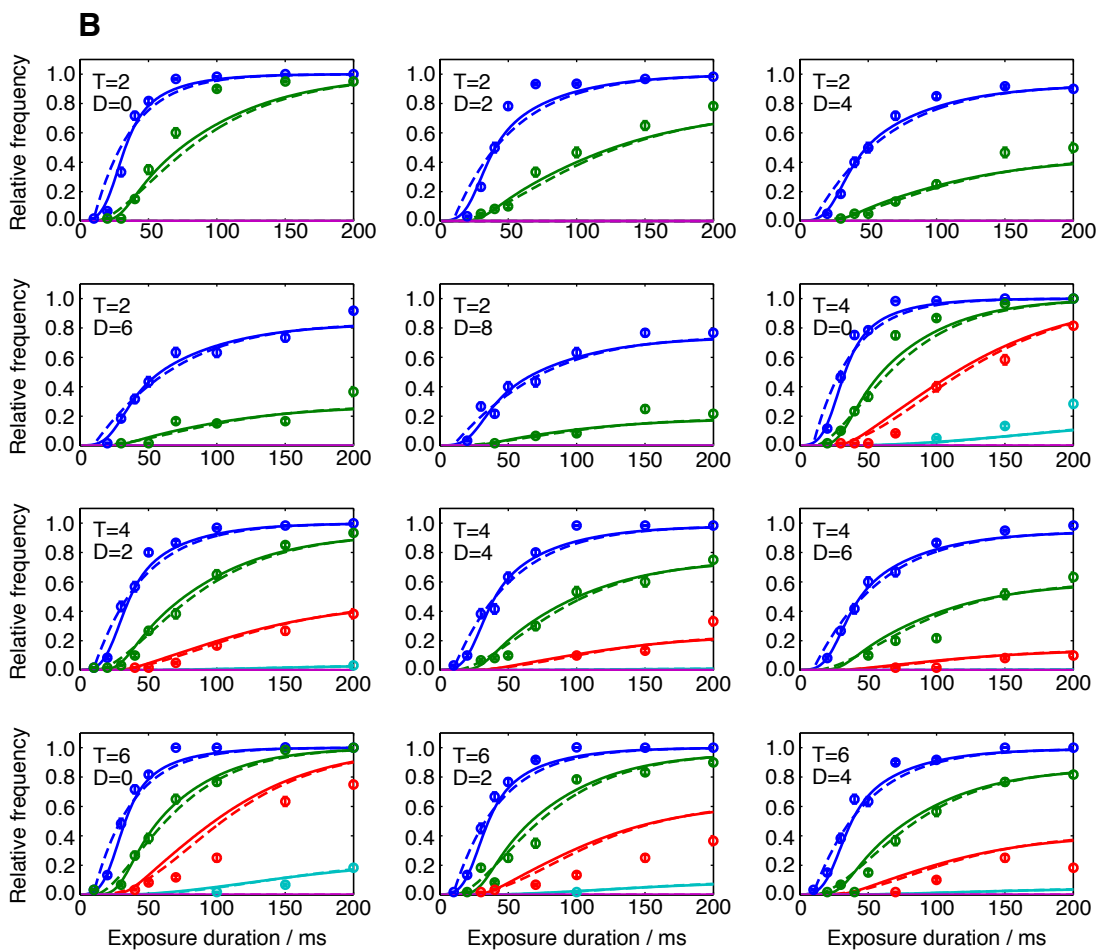
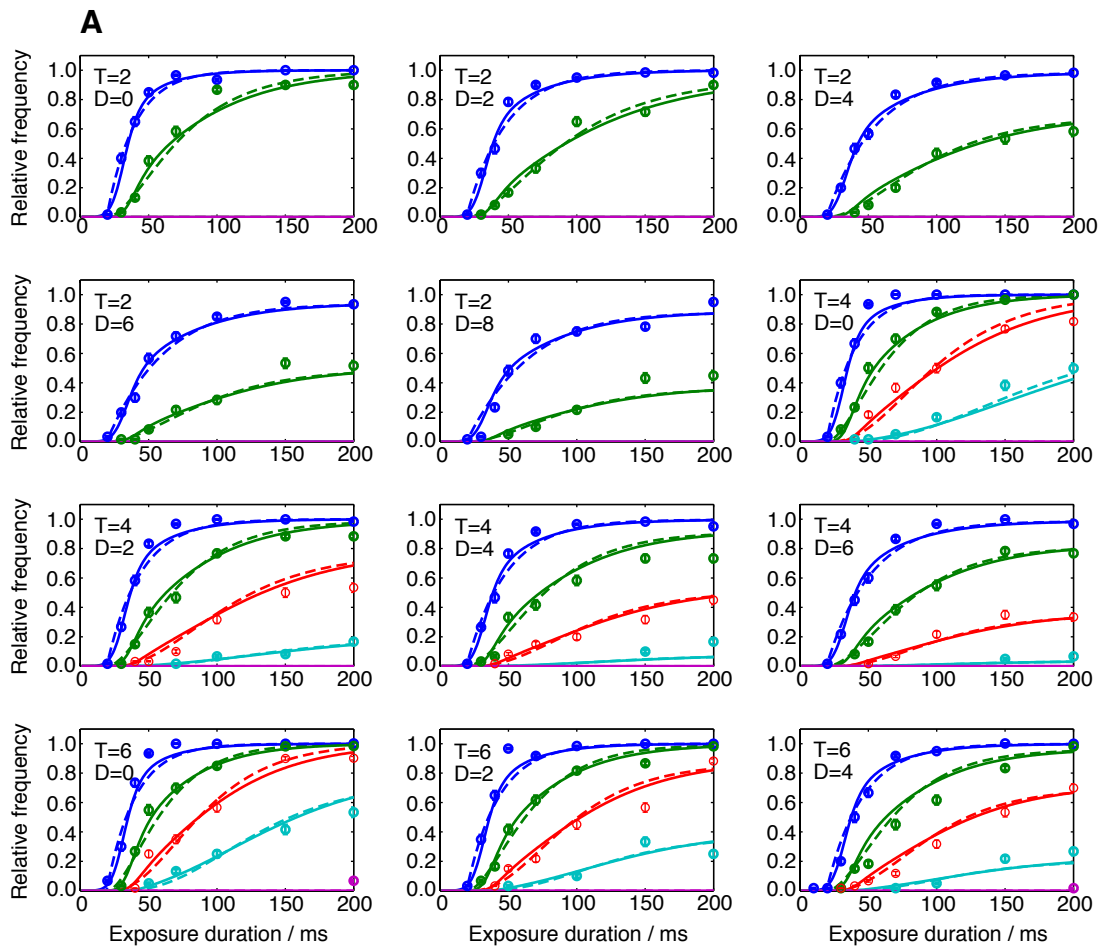
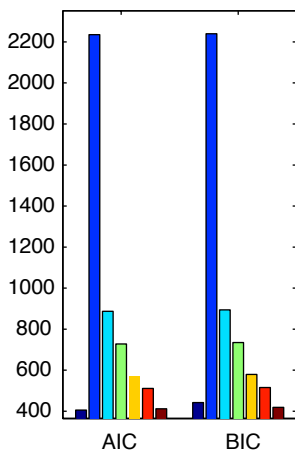
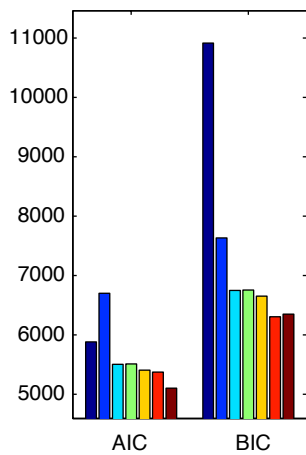


Figure 7

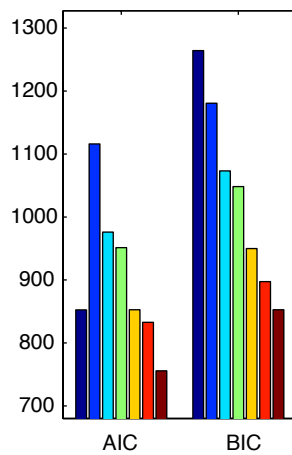
A Experiment 1
(average letter)



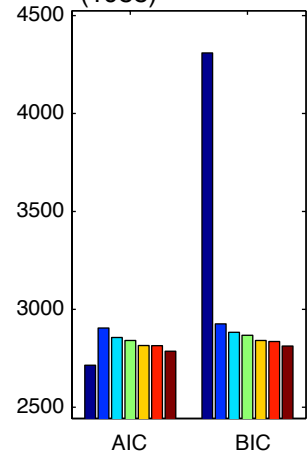
B Experiment 1
(individual letters)



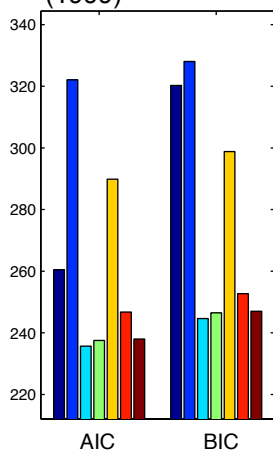
C Experiment 2



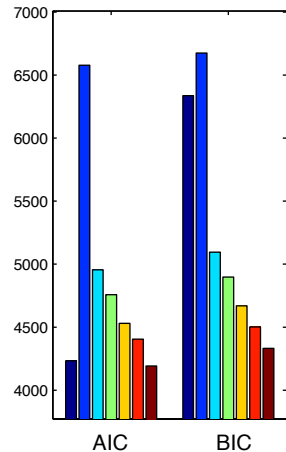
D Whole/partial report
Shibuya & Bundesen
(1988)



E Bundesen & Harms
(1999)



F All experiments



Legend

- Saturated model
- Exponential
- Weibull
- Gamma
- Ex-gaussian
- Log-logistic
- Squared-logistic

References

- Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transactions On Automatic Control*, *AC-19*(6), 716–723.
- Albrecht, D. G., Geisler, W. S., Frazor, R. A., & Crane, A. M. (2002). Visual cortex neurons of monkeys and cats: Temporal dynamics of the contrast response function. *Journal Of Neurophysiology*, *88*(2), 888–913.
- Alexander, K. R., Xie, W., & Derlacki, D. J. (1997). Visual acuity and contrast sensitivity for individual Sloan letters. *Vision Research*, *37*(6), 813–819.
- Alvarez, G., & 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.
- Alzheimer's Association. (2009). 2009 Alzheimer's disease facts and figures. *Alzheimer's and Dementia*, *5*(3), 234-270. doi: 10.1016/j.jalz.2009.03.001
- Baddeley, A. D., Baddeley, H. A., Bucks, R. S., & Wilcock, G. K. (2001). Attentional control in Alzheimer's disease. *Brain*, *124*(8), 1492-1508. doi: 10.1093/brain/124.8.1492
- Baddeley, A. D., & Scott, D. (1971). Short-term forgetting in the absence of proactive interference. *The Quarterly Journal of Experimental Psychology*, *23*(3), 275-283.
- Beeck, H. O. D., Wagemans, J., & Vogels, R. (2007). Effects of perceptual learning in visual backward masking on the responses of macaque inferior temporal neurons. *Neuroscience*, *145*(2), 775-789. doi: 10.1016/j.neuroscience.2006.12.058

- Bjerrum, J. (1882). *Undersøgelser over Formsans og Lyssans i forskellige Øjensygdomme (Investigations on the form sense and light sense in various eye diseases)* (Doctoral Dissertation). University of Copenhagen.
- Bjerrum, J. (1889). Bemærkninger om Formindsnelsen af Synsstyrken (An addendum to the usual examination of the visual field of glaucoma). *Nordisk ophtalmologisk Tidsskrift*, *1*(2), 95.
- Brown, J. (1958). Some tests of the decay theory of immediate memory. *The Quarterly Journal of Experimental Psychology*, *10*(1), 12-21.
- Brüederl, J., & Diekmann, A. (1995). The Log-Logistic Rate Model - 2 Generalizations With An Application To Demographic-Data. *Sociological Methods & Research*, *24*(2), 158–186.
- Bublak, P., Redel, P., & Finke, K. (2006). Spatial and non-spatial attention deficits in neurodegenerative diseases: Assessment based on Bundesen's theory of visual attention (TVA). *Restorative Neurology and Neuroscience*, *24*(4-6), 287-301.
- Bublak, P., Redel, P., Sorg, C., Kurz, A., Förstl, H., Müller, H. J., et al. (2009). Staged decline of visual processing capacity in mild cognitive impairment and Alzheimer's disease. *Neurobiology of Aging*. doi: 10.1016/j.neurobiolaging.2009.07.012
- Bundesen, C. (1990). A Theory of Visual Attention. *Psychological Review*, *97*(4), 523–547.
- Bundesen, C. (1993). The Relationship Between Independent Race Models And Luce Choice Axiom. *Journal Of Mathematical Psychology*, *37*(3), 446–471.
- Bundesen, C. (1998). A computational theory of visual attention. *Philosophical Transactions Of The Royal Society Of London Series B-Biological Sciences*, *353*(1373), 1271–1281.

-
- Bundesen, C., & Habekost, T. (2008). *Principles of Visual Attention: Linking Mind and Brain* (1st ed.). Oxford University Press, USA.
- Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2005). A Neural Theory of Visual Attention: Bridging Cognition and Neurophysiology. *Psychological Review*, *112*(2), 291–328.
- Bundesen, C., & Harms, L. (1999). Single-letter recognition as a function of exposure duration. *Psychological Research-Psychologische Forschung*, *62*(4), 275–279.
- Cattell, J. M. (1886). The inertia of the eye and brain. *Brain*, (8), 295-312.
- Cowan, N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87-185.
- Cronin-Golomb, A., Corkin, S., & Growdon, J. H. (1995). Visual dysfunction predicts cognitive deficits in Alzheimer's disease. *Optometry and Vision Science: Official Publication of the American Academy of Optometry*, *72*(3), 168-176.
- Cronin-Golomb, A., Gilmore, G. C., Nearing, S., Morrison, S. R., & Laudate, T. M. (2007). Enhanced stimulus strength improves visual cognition in aging and Alzheimer's disease. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *43*(7), 952-966.
- Cronin-Golomb, A., & Hof, P. (2004). *Vision in Alzheimer's Disease*. Basel: KARGER. Retrieved from <http://content.karger.com/ProdukteDB/produkte.asp?Doi=79997>
- Deco, G., & Rolls, E. T. (2004). A neurodynamical cortical model of visual attention and invariant object recognition. *Vision Research*, *44*(6), 621-642.
- Deco, G., & Zihl, J. (2004). A biased competition based neurodynamical model of visual neglect. *Medical Engineering & Physics*, *26*(9), 733-743. doi: 10.1016/j.medengphy.2004.06.011
- Diekmann, A. (1992). The Log-Logistic Distribution As A Model For Social Diffusion-

- Processes. *Journal Of Scientific & Industrial Research*, 51(3), 285–290.
- Ditlevsen, S. (2007). A result on the first-passage time of an Ornstein-Uhlenbeck process. *Statistics & Probability Letters*, 77(18), 1744–1749.
- Dreyer, V., Edmund, J., & Møller, P. M. (1992). The first Danish chairs of ophthalmology. *Documenta Ophthalmologica. Advances in Ophthalmology*, 81(1), 87-96.
- 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.
- Duncan, J., Bundesen, C., Olson, A., Humphreys, G., Ward, R., Kyllingsbæk, S., et al. (2003). Attentional Functions in Dorsal and Ventral Simultanagnosia. *Cognitive Neuropsychology*, 20, 675-701. doi: 10.1080/02643290342000041
- Efron, B., & Tibshirani, R. (1994). *An Introduction to the Bootstrap* (1st ed.). Chapman & Hall/CRC, USA.
- Fiset, D., Blais, C., Arguin, M., Tadros, K., Ethier-Majcher, C., Bub, D., et al. (2008a). The spatio-temporal dynamics of visual letter recognition. *Cogn Neuropsychol*, 1. doi: 10.1080/02643290802421160
- Fiset, D., Blais, C., Arguin, M., et al. (2008b). The spatio-temporal dynamics of visual letter recognition. *Cogn Neuropsychol*, 1. doi: 10.1080/02643290802421160
- Fiset, D., Blais, C., Ethier-Majcher, C., Arguin, M., Bub, D., & Gosselin, F. (2008a). Features for identification of uppercase and lowercase letters. *Psychol Sci*, 19(11), 1161–1168. doi: 10.1111/j.1467-9280.2008.02218.x
- Fiset, D., Blais, C., Ethier-Majcher, C., Arguin, M., Bub, D., & Gosselin, F. (2008b). Features for identification of uppercase and lowercase letters. *Psychol Sci*, 19(11), 1161–1168. doi: 10.1111/j.1467-9280.2008.02218.x

-
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189-198.
- Geisler, W. S. (2003). Ideal Observer analysis. In L. Chalupa and J. Werner (Eds.), *The Visual Neurosciences* (pp. 825-837). Boston: MIT press.
- Gervais, M. J., Harvey, L. O., & Roberts, J. O. (1984). Identification Confusions Among Letters Of The Alphabet. *Journal Of Experimental Psychology-Human Perception And Performance*, *10*(5), 655–666.
- Ghosh, J. K., Delampady, M., & Samanta, T. (2006). *An introduction to Bayesian analysis*. Springer.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: John Wiley and Sons.
- 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. doi: 10.1016/S0028-3932(03)00018-6
- Habekost, T., & Starrfelt, R. (2008). Visual attention capacity: A review of TVA-based patient studies. *Scand J Psychol*, *50*(1), 23–32. doi: 10.1111/j.1467-9450.2008.00681.x
- Karlin, S., & Taylor, H. M. (1975). *A First Course in Stochastic Processes, Second Edition* (2nd ed.). Academic Press.
- Lansky, P., Pokora, O., & Rospars, J. P. (2007). Stimulus-response curves in sensory neurons: how to find the stimulus measurable with the highest precision. In *Advances in Brain, Vision, and Artificial Intelligence. Proceedings Second International Symposium* (pp. 338-349). Presented at the BVAI 2007, Naples, Italy.

- Lansky, P., & Ditlevsen, S. (2008). A review of the methods for signal estimation in stochastic diffusion leaky integrate-and-fire neuronal models. *Biol. Cybern.*, *99*(4-5), 253-262.
- Lansky, P., Pokora, O., & Rospars, J. (2008). Classification of stimuli based on stimulus-response curves and their variability. *Brain Res*, *1225*, 57–66. doi: 10.1016/j.brainres.2008.04.058
- Lee, M. L. T., & Whitmore, G. A. (2006). Threshold regression for survival analysis: Modeling event times by a stochastic process reaching a boundary. *Statistical Science*, *21*(4), 501–513.
- Liu, L., Klein, S. A., Xue, F., Zhang, J., & Yu, C. (2009). Using geometric moments to explain human letter recognition near the acuity limit. *J. Vis.*, *9*(1), 1-18.
- Luce, R. D. (1991). *Response Times: Their Role in Inferring Elementary Mental Organization*. Oxford University Press, USA.
- Luck, S., & Vogel, E. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279-281.
- Maloney, L. T., Dal Martello, M. F., Sahn, C., & Spillmann, L. (2005). Past trials influence perception of ambiguous motion quartets through pattern completion. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(8), 3164-3169. doi: 10.1073/pnas.0407157102
- Mattes, J. A. (1997). Can the sensitivity of the Alzheimer's Disease Assessment Scale be increased? *The American Journal of Geriatric Psychiatry: Official Journal of the American Association for Geriatric Psychiatry*, *5*(3), 258-260.
- Mattsson, N., Zetterberg, H., Hansson, O., Andreasen, N., Parnetti, L., Jonsson, M., et al. (2009). CSF biomarkers and incipient Alzheimer disease in patients with mild cognitive impairment. *JAMA: The Journal of the American Medical*

-
- Association*, 302(4), 385-393. doi: 10.1001/jama.2009.1064
- Mendola, J. D., Cronin-Golomb, A., Corkin, S., & Growdon, J. H. (1995). Prevalence of visual deficits in Alzheimer's disease. *Optometry and Vision Science: Official Publication of the American Academy of Optometry*, 72(3), 155-167.
- 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.
- Mozer, M. C., Kinoshita, S., & Shettel, M. (2007). Sequential dependencies offer insight into cognitive control. In *W. Gray (Ed.), Integrated Models of Cognitive Systems* (pp. 180-193). Oxford University Press, USA.
- Oruç, I., Landy, M. S., & Pelli, D. G. (2006). Noise masking reveals channels for second-order letters. *Vision Res*, 46(8-9), 1493–1506. doi: 10.1016/j.visres.2005.08.016
- Palmer, J., Huk, A. C., & Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of Vision*, 5(5), 376-404. doi: 10.1167/5.5.1
- Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research*, 40(10-12), 1227-1268.
- Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *J Opt Soc Am A*, 2(9), 1508–1532.
- Pelli, D. G. (1987). On the relation between summation and facilitation. *Vision Res*, 27(1), 119–123.
- Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research*, 46(28), 4646–4674. doi: 10.1016/j.visres.2006.04.023
- Petersen, A., & Andersen, T. S. (2009). The effect of exposure duration on visual

- character identification in single, whole and partial report, (*submitted*).
- Petersen, A., Kyllingsbæk, S., & Hansen, L. K. (2009). Towards a neural network model of the visual short-term memory. In *In A. Howes, D. Peebles, R. Cooper (Eds.)*, Presented at the 9th International Conference on Cognitive Modeling - ICCM2009, Manchester, UK.
- Petersen, R. C., & Trojanowski, J. Q. (2009). Use of Alzheimer Disease Biomarkers: Potentially Yes for Clinical Trials but Not Yet for Clinical Practice. *JAMA*, 302(4), 436-437. doi: 10.1001/jama.2009.1073
- Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology*, 58, 193-198.
- Phillips, W. A., & Baddeley, A. D. (1971). Reaction-time and short-term memory. *Psychonomic Science*, 22(2), 73-74.
- Ricciardi, L. M., & Sato, S. (1988). 1st-Passage-Time Density And Moments Of The Ornstein-Uhlenbeck Process. *Journal Of Applied Probability*, 25(1), 43-57.
- Ricciardi, L. M., & Sacerdote, L. (1979). The Ornstein-Uhlenbeck process as a model for neuronal activity. I. Mean and variance of the firing time. *Biological Cybernetics*, 35(1), 1-9.
- Scharnowski, F., Hermens, F., & Herzog, M. H. (2007). Bloch's law and the dynamics of feature fusion. *Vision Research*, 47(18), 2444-2452. doi: 10.1016/j.visres.2007.05.004
- Schwartz, S. H. (1998). *Visual Perception: A Clinical Orientation* (2nd ed.). Appleton & Lange.
- Schwarz, G. (1978). Estimating Dimension Of A Model. *Annals Of Statistics*, 6(2), 461-464.
- Shibuya, H. (1994, September 10). *Time course of elementary processes underlying*

-
- visual selective attention: Shape of hazard functions*. Lisbon, Portugal.
- Shibuya, H., & Bundesen, C. (1988). Visual Selection from Multielement Displays: Measuring and Modeling Effects of Exposure Duration. *Journal of Experimental Psychology Human Perception Performance*, *14*(4), 591–600.
- Singpurwalla, N. D. (1995). Survival In Dynamic Environments. *Statistical Science*, *10*(1), 86–103.
- Smith, P. L. (2000). Stochastic Dynamic Models of Response Time and Accuracy: A Foundational Primer*1. *Journal of Mathematical Psychology*, *44*(3), 408-463.
doi: 10.1006/jmps.1999.1260
- Smith, P. L. (1998). Bloch's law predictions from diffusion process models of detection. *Australian Journal Of Psychology*, *50*(3), 139-147.
- Snellen, H. (1862). *Optotypi ad visum determinandum (Probebuchstaben zur Bestimmung der Sehschärfe)*. Utrecht: van de Weijer.
- Sperling, G. (1960). The Information Available in Brief Visual Presentations. *Psychological Monographs*, *74*(11), 1-29.
- Strasburger, H. (2001). Invariance of the psychometric function for character recognition across the visual field. *Perception & Psychophysics*, *63*(8), 1356–1376.
- Strasburger, H. (2005). Character recognition and Ricco's law [Abstract]. *Journal of Vision*, *5*(8), 219a. doi: 10.1167/5.8.219
- Townsend, J. T. (1971). Theoretical Analysis Of An Alphabetic Confusion Matrix. *Perception & Psychophysics*, *9*(1A), 40–&.
- Treutwein, B., & Strasburger, H. (1999). Fitting the psychometric function. *Perception & Psychophysics*, *61*(1), 87–106.
- Usher, M., & Cohen, J. (1999). Short Term Memory and Selection Processes in a

- Frontal-Lobe Model. In *Connectionist Models in Cognitive Neuroscience* (pp. 78-91). Presented at the 5th Neural Computation and Psychology Workshop, Birmingham: Springer-Verlag.
- Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychol Rev*, *108*(3), 550–592.
- Van Breukelen, G. J. P. (1995). Psychometric And Information-Processing Properties Of Selected Response-Time Models. *Psychometrika*, *60*(1), 95–113.
- Van Zandt, T. (2002). Steven's Handbook of Experimental Psychology. In S. S. Stevens & H. E. Pashler (Eds.), (3rd ed.). John Wiley and Sons, USA.
- Whitmore, G. A. (1986). First-passage-time models for duration data: regression structures and competing risks, *35*, 207-219.
- Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception & Psychophysics*, *63*(8), 1314–1329.
- Wolford, G., Marchak, F., & Hughes, H. (1988). Practice effects in backward masking. *Journal of Experimental Psychology. Human Perception and Performance*, *14*(1), 101-112.
- Wouters, H., van Gool, W. A., Schmand, B., & Lindeboom, R. (2008). Revising the ADAS-cog for a More Accurate Assessment of Cognitive Impairment. *Alzheimer Disease & Associated Disorders*, *22*(3), 236-244. doi: 10.1097/WAD.0b013e318174f8b9
- Aalen, O. O., Borgan, Ø., & Gjessing, H. K. (2008). First passage time models: Understanding the shape of the hazard rate. In *Survival and Event History Analysis A Process Point of View* (pp. 387-424). New York: Springer.
- Aalen, O. O., & Gjessing, H. K. (2001). Understanding the shape of the hazard rate: A

process point of view. *Statistical Science*, 16(1), 1–14.

Aalen, O. O., & Gjessing, H. K. (2004). Survival models based on the Ornstein-Uhlenbeck process. *Lifetime Data Analysis*, 10(4), 407–423.