

Université de Montréal

Redundancy gain: manifestations, causes and predictions

par

Sonja Engmann

Département de Psychologie
Faculté des Arts et des Sciences

Thèse présentée à la Faculté des études supérieures et postdoctorales
en vue de l'obtention du grade de Ph.D.
en Psychologie
option Sciences Cognitives et Neuropsychologie

Avril 2009

© Sonja Engmann

Université de Montréal
Faculté des études supérieures et postdoctorales

Cette thèse intitulée :

Redundancy gain: manifestations, causes and predictions

présentée par :

Sonja Engmann

a été évaluée par un jury composé des personnes suivantes :

Martin Arguin
Président-rapporteur

Denis Cousineau
Directeur de recherche

Pierre Jolicoeur
Membre du jury

Aaron Johnson
Examineur externe

Stéphane Molotchnikoff
Représentant du doyen

ABSTRACT

Keywords: redundancy gain, coactivation, analysis of response time distributions

Response times in a visual object recognition task decrease significantly if targets can be distinguished by two redundant attributes. Redundancy gain for two attributes is a common finding, but redundancy gain from three attributes has been found only for stimuli from three different modalities (tactile, auditory, and visual). This study extends those results by showing that redundancy gain from three attributes within the visual modality is possible. It also provides a more detailed investigation of the characteristics of redundancy gain. Apart from a decrease in response times for redundant targets, these include a decrease in minimal response times and an increase in symmetry of the response time distribution.

This study further presents evidence that neither race models nor coactivation models can account for all characteristics of redundancy gain. In this context, we discuss the problem of calculating an upper limit for the performance of race models for triple redundant targets, and introduce a new method of evaluating triple redundancy gain based on performance for double redundant targets. In order to explain the results from this study, the cascade race model is introduced. The cascade race model consists of several input channels, which are triggered by a cascade of activations before satisfying a single decision criterion,

and is able to provide a unifying approach to previous research on the causes of redundancy gain.

The analysis of the characteristics of response time distributions, including their mean, symmetry, onset, and scale, is an essential tool in this study. It was therefore important to find an adequate statistical test capable of reflecting differences in all these characteristics. We discuss the problem and importance of analysing response times without data loss, as well as the inadequacy of common methods of analysis such as the pooling of response times across participants (e.g. Vincentizing) in the present context.

We present tests of distributions as an alternative method for comparing distributions, response time distributions in particular, the most common of these being the Kolmogorov-Smirnoff test. We also introduce a test yet unknown in psychology: the two-sample Anderson-Darling test of goodness of fit. We compare both tests, presenting conclusive evidence that the Anderson-Darling test is more accurate and powerful: when comparing two distributions that vary (1) in onset only, (2) in scale only, (3) in symmetry only, or (4) that have the same mean and standard deviation but differ on the tail ends only, the Anderson-Darling test proves to detect differences better than the Kolmogorov-Smirnoff test. Finally, the Anderson-Darling test has a type I error rate corresponding to *alpha* whereas the Kolmogorov-Smirnoff test is overly conservative. Consequently, the Anderson-Darling test requires less data than the Kolmogorov-Smirnoff test to reach sufficient statistical power.

RÉSUMÉ

Mots-clés: gain de redondance, coactivation, analyse des distributions de temps de réponse

Les temps de réponse dans une tâche de reconnaissance d'objets visuels diminuent de façon significative lorsque les cibles peuvent être distinguées à partir de deux attributs redondants. Le gain de redondance pour deux attributs est un résultat commun dans la littérature, mais un gain causé par trois attributs redondants n'a été observé que lorsque ces trois attributs venaient de trois modalités différentes (tactile, auditive et visuelle). La présente étude démontre que le gain de redondance pour trois attributs de la même modalité est effectivement possible. Elle inclut aussi une investigation plus détaillée des caractéristiques du gain de redondance. Celles-ci incluent, outre la diminution des temps de réponse, une diminution des temps de réponses *minimaux* particulièrement et une augmentation de la symétrie de la distribution des temps de réponse.

Cette étude présente des indices que ni les modèles de course, ni les modèles de coactivation ne sont en mesure d'expliquer l'ensemble des caractéristiques du gain de redondance. Dans ce contexte, nous introduisons une nouvelle méthode pour évaluer le triple gain de redondance basée sur la performance des cibles doublement redondantes. Le modèle de cascade est présenté afin d'expliquer les résultats de cette étude. Ce modèle comporte plusieurs voies de traitement qui sont déclenchées par une cascade d'activations

avant de satisfaire un seul critère de décision. Il offre une approche homogène aux recherches antérieures sur le gain de redondance.

L'analyse des caractéristiques des distributions de temps de réponse, soit leur moyenne, leur symétrie, leur décalage ou leur étendue, est un outil essentiel pour cette étude. Il était important de trouver un test statistique capable de refléter les différences au niveau de toutes ces caractéristiques. Nous abordons la problématique d'analyser les temps de réponse sans perte d'information, ainsi que l'insuffisance des méthodes d'analyse communes dans ce contexte, comme grouper les temps de réponses de plusieurs participants (e. g. Vincentizing).

Les tests de distributions, le plus connu étant le test de Kolmogorov-Smirnoff, constituent une meilleure alternative pour comparer des distributions, celles des temps de réponse en particulier. Un test encore inconnu en psychologie est introduit : le test d'Anderson-Darling à deux échantillons. Les deux tests sont comparés, et puis nous présentons des indices concluants démontrant la puissance du test d'Anderson-Darling : en comparant des distributions qui varient seulement au niveau de (1) leur décalage, (2) leur étendue, (3) leur symétrie, ou (4) leurs extrémités, nous pouvons affirmer que le test d'Anderson-Darling reconnaît mieux les différences. De plus, le test d'Anderson-Darling a un taux d'erreur de type I qui correspond exactement à l'*alpha* tandis que le test de Kolmogorov-Smirnoff est trop conservateur. En conséquence, le test d'Anderson-Darling nécessite moins de données pour atteindre une puissance statistique suffisante.

CONTENTS

ABSTRACT	iii
RÉSUMÉ	v
CONTENTS	vii
List of tables	xii
List of figures	xiii
List of abbreviations	xv
Dedication	xvi
Acknowledgements	xvii
CHAPTER 1: Introduction	1
Analysis	6
Threshold of race model performance	6
<i>Miller Inequality</i>	6
<i>Townsend Bound</i>	8
Differentiating between response time distributions	11
Empirical data	15
Pilot experiment A (3RedA)	15
<i>Method</i>	15
<i>Results</i>	20
Further research: General method	25
Practice effects	27
<i>3RedB : Colour, form and letter</i>	27

Effects of masking.....	32
<i>3RedC : Masked</i>	32
<i>2RedM&M : Masked / Unmasked</i>	36
Early processing stages.....	39
<i>3RedD : colour, orientation, frequency</i>	40
<i>3RedDSat : low saturation</i>	44
<i>3RedE1 : colour, form, and movement</i>	45
Do certain brain areas favour coactivation?.....	46
<i>2RedMST : perceived depth and direction of movement</i>	47
Characteristics of redundancy gain.....	53
Modelling.....	55
Final research questions.....	57
CHAPTER 2: Coactivation results cannot be explained by pure coactivation	
models	58
Déclaration des coauteurs.....	59
Apport original.....	60
Abstract.....	61
Introduction.....	62
Method.....	71
Participants.....	71
Stimuli and Apparatus.....	71
Design.....	72

Procedure	73
Results	75
Difference in negative contingencies	76
Redundancy gain	77
Excluding Race Models.....	80
Excluding race models for gain from double to triple redundancy	83
Characteristics of the RT distributions	86
Predictions of the coactivation model	89
The cascade race model.....	92
The serial coactivation model.....	94
The parallel coactivation model	96
Some conclusions related to model fitting	100
Discussion	102
Acknowledgements	109
References	110
Appendix A: Simulating the coactivation model	120
Tables	122
Figures	127
CHAPTER 3: Comparing distributions: the two-sample Anderson-Darling test	
as an alternative to the Kolmogorov-Smirnoff test.....	138
Déclaration des coauteurs.....	139
Apport original	140

Abstract	141
Introduction	142
Why combining response time distributions across participants is problematic	143
Different methods of comparing distributions participant-by-participant..	144
Comparison of Kolmogorov-Smirnoff and Anderson-Darling tests	148
Kolmogorov-Smirnoff Test	148
Anderson-Darling Test	150
Comparison of the two tests when shift, scale and symmetry are varied independently	153
<i>Method</i>	154
<i>Results</i>	155
Comparison of the two tests when D_1 and D_2 differ in the tails only	158
<i>Method</i>	158
<i>Results</i>	159
Sample size needed to reach sufficient statistical power when shift, scale and symmetry are varied independently	159
<i>Method</i>	161
<i>Results</i>	161
Discussion	164
Acknowledgments	166
References	167
Appendix	172

Tables	174
Figures	177
CHAPTER 4: Conclusion	183
GENERAL REFERENCES	186
APPENDIX	195
Instructions for participants	195

List of tables

CHAPTER 1

<i>Table 1.</i> Stimulus distribution for one block for three target attributes.....	19
--	----

CHAPTER 2

<i>Table 1.</i> Stimulus distribution for one block of 60 trials	122
--	-----

<i>Table 2.</i> Proportion of participants for whom the distribution in a condition was significantly different from the distribution in a condition with less redundancy	123
---	-----

<i>Table 3.</i> Summary of the various predictions made by race models.....	124
---	-----

<i>Table 4.</i> Fit of the models to the response time distributions of the correct response	125
---	-----

<i>Table 5.</i> Mean fit and mean parameter values for the models tested	126
--	-----

CHAPTER 3

<i>Table 1.</i> Parameters of the Weibull and Normal distributions.....	174
---	-----

<i>Table 2.</i> Definition of large, medium and small effect size for the three parameters of the Weibull distribution.....	175
--	-----

<i>Table 3.</i> Number of data required to reach a power of 80%	176
---	-----

List of figures

CHAPTER 1

Figure 1. Examples of stimuli.....	18
Figure 2. Experiment 3RedA.....	22
Figure 3. Experiment 3RedB.....	30
Figure 4. Experiment 3RedC.....	35
Figure 5. 2RedM&M: effect of condition and masking on response times.....	38
Figure 6. Experiment 3RedD.....	43
Figure 7. Examples of stimuli for the experiment 2RedMST.....	49
Figure 8. Experiment 2RedMST.....	52

CHAPTER 2

Figure 1. Mean response time per condition.....	127
Figure 2. RT distributions of participant 13 for single- and double redundant conditions.....	128
Figure 3. RT distributions of participant 13 for double and triple redundant conditions.....	129
Figure 4. Townsend Bound and RT distributions of participant 13 for double redundant conditions.....	130
Figure 5. Townsend Bound and RT distributions of participant 13 for triple redundant conditions.....	131
Figure 6. Townsend Bound and Townsend Bound calculated from double redundant distributions.....	132

Figure 7. Parameters of RT distributions.....	133
Figure 8. Parameters of RT distributions – real and simulated	134
Figure 9. Different versions of a race model with three possible inputs ...	136
Figure 10. Influence of the parameters ρ (internal redundancy), R (external redundancy) and k (decision threshold) on the observed symmetry.....	137

CHAPTER 3

Figure 1. The proportion of significant differences between the two distributions for the AD and KS test as a function of $\Delta 1$ (changes in shift).....	177
Figure 2. Absolute advantage of AD over KS test as a function of $\Delta 1$ (changes in shift).....	178
Figure 3. The proportion of significant differences between the two distributions for the AD and KS test as a function of $\Delta 2$ (changes in scale)	179
Figure 4: The proportion of significant differences between the two distributions for the AD and KS test as a function of $\Delta 3$ (changes in asymmetry)	180
Figure 5. Weibull and Normal distributions used for evaluation of performance when distributions differ at tails	181
Figure 6. The two distributions compared when the effect size of change in symmetry is large (left) and small (right)	182

List of abbreviations

AD test	Anderson-Darling test
°VA	degree of visual angle
KS test	Kolmogorov-Smirnoff test
ms	milliseconds
P	probability
RT	response times
RTE	Redundant Target Effect
std	standard deviation

Für meine Eltern

Acknowledgements

First and foremost, I would like to thank my supervisor, Denis Cousineau. His support - professional, financial and moral - through experimental and statistical problems, visa troubles and loss of motivation, has been greatly appreciated, and this thesis wouldn't have been possible without it.

My colleagues, especially Jade Girard and Sophie Callies, have made my studies so much more enjoyable and fun. I would like to thank them for the numerous animated discussions, lunch breaks in company of M Hubertu, and all other moments of procrastination we shared.

The encouragement and support of my friends has been invaluable. I want to thank in particular: The Rhizome, for its warmth, colourfulness and craziness, for being my home, for providing the balance to sitting in front of a computer; the Coop sur Genereux, for feeding me and letting me crash; Zbyněk, for Hostyn, for always being there, and listening patiently to the latest developments in my thesis.

And finally, I am very grateful for the continued support and love of my parents and brother. They have given me the freedom and strength to become the person I am today. And they have been very patiently waiting for me to finally finish my studies!

CHAPTER 1

INTRODUCTION

Recognizing objects is a seemingly simple, even trivial task to ask of a human. However, when trying to disassemble the process of object recognition into its different components, or trying to simulate human performance on object recognition tasks, we quickly realize that it is far more complex, and that we are still far from understanding why humans perform so well, and far from achieving close to human performance in models of object recognition.

Visual input is decomposed into its smallest parts (single neuron receptive fields) by the visual processing system. Visual processing is organised in a hierarchical manner, with a series of subsequent processing areas analysing more and more complex combinations of information (Goodale & Milner, 1992; Kandel, Schwarz & Jessell, 2000). Complex objects can either be processed holistically (such as faces; Desimone, 1991; Farah, 1990) or analytically, that is by analysing the constituent parts (Farah, 1990; Biederman, 1990). However, we do not know how complex objects are reconstituted from their individual components. How does the visual system know which features belong together? This problem is referred to as the binding problem (Treisman & Gelade, 1980, Treisman, 1996). Several possible solutions for the binding problem have been proposed including the 'grandmother cell theory', which postulates highly specialised cells that respond to a specific combination of attributes (Barlow, 1972), spatial proximity

(Wolfe, Cave & Franzel, 1989), and the theory of synchronized firing of cells responding to the same object (Milner, 1974; Singer & Gray, 1995). Although the debate is not resolved yet, a combination of combined selectivity, synchronized firing and spatial proximity seems most likely to account for binding of objects (Treisman, 1996).

Spatial proximity, perceived continuity, or similarity of features can account for what is perceived as an object (Treisman, 1990; Palmer, 1981; Gestalt Psychology: Köhler, 1947, among others). This is generally known as an effect of grouping. Contrast also plays a major role in object recognition. It is the most important and most studied tool for defining what constitutes an object, where the edges of an object are (Marr, 1976), and what belongs to another object or background (Lamme, 1995). Our visual system is based upon an analysis of contrast (e.g. luminance contrast, colour, motion or orientation contrast; Livingstone & Hubel, 1988), and contrast plays an important role in the attraction of attention (Engmann et al., in press). Contrast is frequently high around the edges of objects (e.g. there is a difference in colour between an object and its background), and it has been shown that high contrast attracts fixation (Tatler, Baddeley & Gilchrist, 2005) and that the visual cortex responds selectively to objects that are separated from the background by elevated high contrast (Zipser, Lamme & Schiller, 1996; Lamme, 1995).

Object recognition can be facilitated or inhibited by a number of factors, such as familiarity or complexity of the object (Logothetis & Sheinberg, 1996),

familiarity of viewing angle (Tarr & Pinker, 1989), or context (Torralba, Murphy, Freeman & Rubin, 2003). Another characteristic that can facilitate or inhibit object recognition involves the number of target attributes by which an object can be recognised. In certain situations, several distinct attributes that indicate the identity of an object can facilitate object recognition, in other situations, object recognition is inhibited by several target attributes. If a target object is defined by a single attribute that separates it from its surroundings, recognition is facilitated independently of the number of surrounding distracters (Pop-out effect; Treisman & Souther, 1985). If a combination of attributes is needed to unambiguously identify a target (i.e. if the target itself is unique, but shares at least one feature with all surrounding distracters) target recognition is inhibited and becomes dependent on the number of distracters (Treisman & Souther, 1985). However, if a target is defined by a combination of attributes, either of which is sufficient to identify the target, recognition is facilitated. This is known as the redundant target effect (RTE; Kinchla, 1974; Miller 1982). Chapter two contains a review of literature on the RTE, the main body of which studies objects defined by two target attributes. Stimuli defined by three target attributes have rarely been studied (Diederich, 1995), and never within a single modality. An investigation of triple redundancy gain in purely visual stimuli is a novel question.

The initial motivation for the present study was to provide evidence for triple redundancy gain in the visual modality. The main goal was to show that within one modality, redundancy gain is not limited to two features, but that each new target feature added has a facilitatory effect on object recognition. We also

wanted to investigate for which feature combinations a triple redundancy gain is possible. A combination of three visual target features that produces a significant gain in reaction time speed over double redundant targets is therefore sought.

The question of a possible generalisation from a “double” RTE to a “triple” RTE is a very important one. An in-depth investigation would shed much light on information processing within a single modality. It has been shown that the RTE is only possible when parallel processing is assumed (Van der Heijden, La Heij & Boer, 1983, Krummenacher, Müller & Heller, 2001). The absence of additional gain from a third target attribute would, for instance, demonstrate a limit of parallel processing in the visual system. Therefore, the first important question is whether enough resources can be made available for a single modality to process three features in parallel quickly enough to enable a redundancy gain from the third attribute.

Another motivation for this study was to clarify the question of possible causes of redundancy gain (again, chapter two contains a more detailed review of different theories). Several types of models have been proposed to explain the RTE, namely the race model (Raab, 1962), the coactivation model (Smith, 1968; Miller 1982; Schwartz, 1989) and crosstalk (Mordkoff & Yantis, 1991). Race models assume independent channels separately accumulating evidence in favour of the specific signal or feature to which they are tuned. Object recognition occurs when one channel has accumulated enough evidence to overcome its decision threshold. Race models allow redundancy gain because more channels improve the

chances of evidence for one feature being accumulated particularly fast. This probability-based explanation of the RTE allows the calculation of an upper limit to performance of race models (Miller 1982). Coactivation models combine evidences from different channels to satisfy a *single* threshold criterion. If several channels exist, the pooling of activation from several still weakly activated channels will be sufficient to overcome the threshold, thereby causing redundancy gain. Crosstalk models are basically race models with connections between channels, allowing benefit from correlations between different features, and thus causing a redundancy gain.

The goal of this study was to provide experimental data that is able to distinguish between three different theories explaining the RTE in the visual system, and thereby to exclude two out of three of these possible explanations, either *a priori*, through the experimental design, or *a posteriori*, by the use of a decision criterion definitely favouring or excluding one theory based on the experimental results. We hypothesized that coactivation is the rule in the visual system, but might not always be directly observable due to different noise levels, or different processing speeds as a function of the visual features involved.

Analysis

The research goals formulated above posed two methodological problems for the analysis of response time data. First, we had to define a threshold of race model performance for triple redundancy. Second, we needed to find an appropriate statistical test for finding differences between response times from different conditions without losing information.

Threshold of race model performance

Miller Inequality

To provide evidence in favour of or against race models as a cause for gain, most recent studies investigating redundancy gain and coactivation models used the Miller Inequality (Miller, 1978),

$$P(RT < t|T_1 \cap T_2) \leq P(RT < t|T_1) + P(RT < t|T_2) \quad (1)$$

where $P(RT < t|T_i)$ is the probability of participants responding faster than time t given a target (T_i) is present on channel i .

It is conceivable that two feature channels are not independent of each other. Since this dependency can stem from any number of possible causes, many of which could be biological and therefore not directly observable, one cannot estimate the degree of dependence between channels. All one knows is that it must

be equal to or greater than zero. If X equals the degree of dependency between two channels,

$$X = P[(RT < t|T_1) \cap (RT < t|T_2)] \quad (2)$$

then X needs to be subtracted from the sum of response time distributions of the separate channels to calculate race model performance for two redundant channels:

$$P(RT < t|T_1 \cap T_2) = P(RT < t|T_1) + P(RT < t|T_2) - X \quad (3)$$

Subtracting X , an unknown positive quantity, from both sides of the equation leaves us with the Miller Inequality (eq. 1), a definite upper limit to the performance of race models with two channels, and a very efficient criterion of exclusion for race models on any task with two redundant targets.

When generalising equation (3) to three channels, we need to factor in dependencies between any two channels

$$X_{ij} = P[(RT < t|T_i) \cap (RT < t|T_j)] \quad (4)$$

and all three channels,

$$Y_{ijk} = P[(RT < t|T_i) \cap (RT < t|T_j) \cap (RT < t|T_k)] \quad (5)$$

which results in the following equation:

$$P(RT < t|T_1 \cap T_2 \cap T_3) = P(RT < t|T_1) + P(RT < t|T_2) + P(RT < t|T_3) - X_{12} - X_{13} - X_{23} + 2 \times Y_{123} \quad (6)$$

The degree of dependency between all three channels, Y_{123} , an unknown positive number, is a subset of each X_{ij} , and therefore, having been subtracted thrice with each X_{ij} , needs to be added twice to the equation again to make it valid. Both types of dependencies are of unknown positive size, so trying to factor them out of equation (6) makes it impossible to determine in which direction the extension of the Miller Inequality would tend.

Diederich and Colonius, in their 2004 study, used an extension of the Miller Inequality to three channels to refute race models in a stimulus detection task with stimuli from three different modalities (auditory, visual, and/ or tactile). Their extension is not valid for this study, however, since it did not account for the unknown factor of dependency between all three channels, and we cannot assume three channels from a single modality to be completely independent of each other.

Townsend Bound

An alternative to the Miller Inequality was proposed by Townsend and Nozawa (1995; a similar bound was proposed by Mordkoff and Yantis, 1991, p. 535). It is based upon survivor functions (one minus the cumulative distribution) of response times instead of cumulative distribution functions. The upper limit to race model performance with more than one channel is given by the survivor function of the product of the survivor functions of each channel:

$$P(\mathbf{RT}_{123} < t) = 1 - \prod_{1 \leq i \leq 3} 1 - P(\mathbf{RT}_i < t) \quad (7)$$

where \mathbf{RT}_{123} is a response time when target attributes from all three channels are present and i indexes the three channels. If the observed response time distribution in a redundant target task is significantly faster than predicted by this boundary, race models as the sole explanation of redundancy gain are rejected. The Townsend Bound can be calculated for any number of channels.

For a pure race model, the Townsend Bound for three channels based on the three survivor functions of the single channels is perfectly valid. However, we need to consider the possibility of a mixed coactivation and race model being able to explain the results. What if a decision about object identity was made coactively by two channels, but the third channel contributes solely on a winner-take-all basis? So the question is: how do we distinguish between the possibility of a mixed model and a pure coactivation model (i.e. a model where responses from all three channels are pooled to satisfy a single decision criterion)? We calculated a Townsend Bound for mixed models (models where two channels interact by coactivation and the third channel contributes only within the range of statistical facilitation) based on the survivor functions of an RT distribution where target attributes are present on two channels plus the RT distribution of the target attribute on the respectively missing channel. This yields three Townsend Bounds (one for each combination of two plus one channels), which we combine into one

single threshold criterion for mixed models by taking the maximum value out of these three criteria at each time point:

$$P(RT_{123} < t) = \max \left[\begin{array}{l} 1 - (1 - P(RT_{12} < t) \times (1 - (RT_3 < t))), \\ 1 - (1 - P(RT_{13} < t) \times (1 - (RT_2 < t))), \\ 1 - (1 - P(RT_{23} < t) \times (1 - (RT_1 < t))) \end{array} \right] \quad (8)$$

This yields the most liberal evaluation of performance if any of the three channels contributes only by statistical facilitation as a third redundant attribute. Hence, exceeding the limit can only be achieved if all three target attributes contribute significantly by coactivation to the amount of redundancy gain. Alternatively, the Townsend Bound for mixed models could have been calculated from the product of the survivor functions of response times in the three possible conditions when target attributes on any two channels were present, analogous to equation (7). However, in this case, the gain contributed by each attribute would be included twice (once in each of the two double redundant conditions it is part of), thereby obtaining an upper limit which would definitely exceed performance of a combination of coactivation for two and statistical facilitation for the third attribute. Equation (8) ensures that each attribute contributes only once, while still ensuring the best possible performance under the assumption that statistical facilitation is responsible for the gain attributed to the third target attribute.

When testing for triple redundancy, we used the simple Townsend Bound (eq. 7) as a default. Should it be violated, we also tested violation of the mixed-model Townsend Bound (eq. 8).

Differentiating between response time distributions

We decided to analyse response time data on a participant by participant basis for various reasons, mainly because grouping of several participants' response time data is invariably accompanied by some loss of information – information about variance and symmetry which is particular to individual participants. This leads to false representations of RT data, such as flattened or bimodal distributions, which tend towards normality, even if the underlying individual response time distributions are not normally distributed. Using a technique such as Vincentizing (Vincent 1912, Rouder & Speckman, 2004) for grouping avoids bias due to loss of variance information. Vincentizing involves grouping RT distributions by quantiles: response times in the first n^{th} percentile of each RT distribution are averaged, and response times in the next n^{th} percentile, and so on. Distributions are “averaged” by taking into account the relative position of each response time, thereby avoiding flattening or bimodality. However, Vincentized distributions still tend towards normality (Thomas & Ross, 1980), whereas normality cannot be assumed for response time distributions (Logan, 1992; Rouder, Lu, Speckman, Sun & Jiang, 2005).

Several authors used multiple t-tests on quantiles (Miller, 1982; Mordkoff & Yantis, 1991, 1993, among others). Quantiles (e. g. the 5th percent quantiles) are computed for each participant in the two conditions whose distributions are to be compared, and then tested for equality using a t-test. This procedure is replicated for all quantiles at given intervals (e. g. the 10th, the 15th, etc. percent). This

method allows an estimate of where RT distributions of all participants differ significantly. It keeps individual participants' data separate, and analyses more than distribution means.

However, we noticed a large between-participant variability in all our experiments (see next section), the difference between participants being at times larger than the actual effect of redundancy gain. In this case, multiple t-tests cannot detect redundancy gain. Additionally, sample size for each t-test is only as large as the number of participants in an experiment; therefore statistical power may not be sufficient, especially if the effect size is not very large to begin with. Finally, the data at one time point are highly correlated with the data at the previous and following time point, influencing the probability of a type I error. We therefore decided upon a participant by participant analysis.

An analysis of redundancy gain on a participant by participant basis has the advantage of keeping all information particular to a participant, while making the effect of redundancy gain directly observable, without having to factor out between-participant variability. The most common methods of comparing response time distributions, a t-test or an ANOVA, were not an option for analysis, since both assume normality, and only analyse differences in mean and variance of samples. Therefore, the best choice for a participant by participant analysis of response times in the present experimental context is a test of distributions, the most well-known being the Kolmogorov-Smirnoff (KS) test (Kolmogorov, 1941; Smirnoff, 1939). The Anderson-Darling (AD) test (Anderson & Darling, 1952), an

alternative similar to the KS test, is mainly used in the field of engineering and not known in psychology at all.

We implemented a two-sample version of the AD test in Matlab (MathWorks Inc., Natick, MA). After comparing the performance of the KS and the AD test on our experimental data, we noticed that the AD test was more sensitive to small differences between response time distributions. Also there is evidence that the one-sample version of the AD test is especially sensitive to the tail ends of distributions (Darling, 1957). If this holds true for the two-sample version as well, the AD test is even better suited to the current context: we hypothesize that minimal response times are more affected by redundancy than other characteristics of response time distributions. We suspect that the two-sample AD test is a powerful tool for comparing response time distributions, and should be more frequently used in the field of cognitive psychology. After implementing the two-sample version, we therefore decided to test power and reliability of the AD test more rigorously. At the same time this gave us an opportunity to review other techniques for analysing response time distributions (see chapter three for details).

In order to determine which test was more powerful and better suited to the present context, we calculated the probability of both the AD test and the KS test to detect differences in shape, symmetry, shift and behaviour at the extrema of samples drawn from theoretical distributions in a series of Monte Carlo simulations. These are four additional ways, apart from mean and variance, to

characterise response time distributions (especially if these aren't normally distributed). We relied on all this additional information in order to differentiate between response time distributions, and needed the test most adapted to detecting differences in all these dimensions. Also, due to the participant by participant analysis, the amount of data per condition was not very large in most of the subsequently described experiments. We therefore also needed to determine which test would yield greater statistical power (Cohen, 1992) given the expected effect size. The Anderson-Darling test proved superior on the detection of differences between distributions as well as for statistical power. Please refer to chapter three for details of method and results of Monte Carlo simulations and the calculation of statistical power.

Empirical data

The following section provides a description of the progression of this research, and the subsequent evolution of the research questions. A series of pilot experiments were necessary in order to develop the paradigm for the last, successful experiment described in chapter two. Each of these pilot experiments provides interesting results, such as empirical evidence that parallel processing is needed to observe a redundancy gain (2RedMST), or the effect of practice (3RedB) and masking (3RedC and 2RedM&M) on the redundancy gain. This section can be skipped without compromising the comprehension of the remaining chapters.

Pilot experiment A (3RedA)

Method

Participants. Participants were 4 female undergraduate students from the Université de Montréal, between 19 and 25 years of age. All had normal or corrected-to-normal vision. Participants were compensated with 8\$ per hour for their participation.

Stimuli and apparatus. We used simple two-dimensional geometrical objects as stimuli. Stimuli were created in the RGB color space, using Matlab (MathWorks Inc., Natick, MA). Stimuli were presented using E-Prime (Psychology Software Tools, Inc., Pittsburgh, PA) on a SVGA monitor (refresh rate: 85 Hz) at a distance of 80 cm from the participants. The stimuli measured 1.5,

2 or 3 °VA (degrees of visual angle); they were either red, green, or blue; and lastly, their form was a circle, a triangle or a square (see Figure 1 for an example of the stimuli used). Stimuli were presented in front of an equiluminant gray background with stimulus luminance at 50% and stimulus saturation at 100% percent. Target stimuli possessed one or more of the following attributes: color red, form of a circle, and medium size (2 °VA). The presence of any single one of these attributes was sufficient to define a given stimulus as a target. Non-target stimuli did not possess any of the target attributes. They were either green or blue, a triangle or a square and large or small in size.

Design. 50% of all stimuli presented to participants were targets. To avoid contingencies between attributes on different channels which would facilitate redundant target recognition, the stimulus distribution shown in Table I (top) is an extension of the distribution suggested by Mordkoff and Yantis (1991), following the three rules of contingency formulated by Mordkoff and Yantis (1991).

Procedure. The experiment consisted of 17 blocks with 42 trials per block for a total of 714 trials. Stimulus distribution did not vary between blocks, but the order of trials was randomized. Participants had the possibility to take a break between blocks. The triple redundant target (target with all three target attributes present), the three double redundant targets (any two target attributes present, plus one of two possible distracters on the third channel), and the three stimuli with only one target attribute present, were presented 51 times per participant. Non-target stimuli were presented 357 times per participant. Each trial started with the

presentation of a fixation point for 850 ms. The stimulus was then presented for 1000 ms. Finally, a feedback slide appeared for 1200 ms, which was followed by a blank screen for 1000 ms in preparation for the next trial.

We used a Go-NoGo experimental paradigm. Participants were required to press the SPACE key on a keyboard as soon as they recognized a target stimulus, and discouraged from doing so if they recognized a non-target. They were encouraged to respond as fast as possible while making as few errors as possible. Responses had to happen within a time frame of 0 to 1000 ms after stimulus onset.

Participants received feedback on their performance on each trial. Feedback on false responses was accompanied by a 700 Hz sound. Fast and correct performance was further encouraged by a system of points: participants were encouraged to try for the best score. Participants received 30 points for hits and 15 for correct rejections, 50 for particularly fast hits (under 300 ms), and -350 points for false alarms and misses. At the end of each block participants were given their cumulative score.

Figure 1. Examples of stimuli. First column shows stimuli with all three target attributes present, second and third columns with three non-target attributes.

- a) Stimuli for experiment 3RedA
- b) Stimuli for experiments 3RedB and 3RedC
- c) Stimuli for experiments 3RedD and 3RedDSat
- d) Stimuli for experiment 3RedE

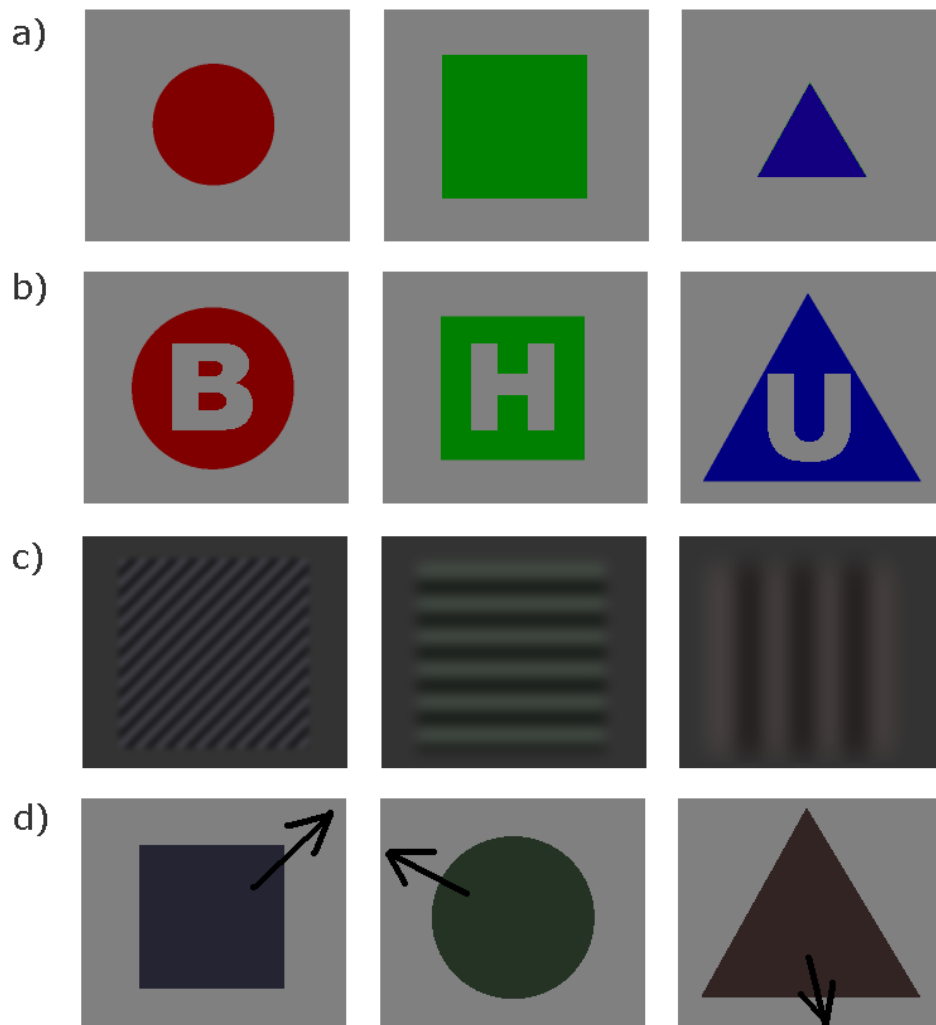


Table I. Stimulus distribution for one block for three target attributes, for 3redA (top) and all subsequent experiments (bottom). Panels, rows and columns show the different attribute values. White fields are non targets, light gray fields are stimuli with one target attribute, medium gray fields are stimuli with two target attributes, and the dark gray field represents a stimulus with all three target attributes present.

Feature A											
Feature value A1 (target)				A2				A3			
Feature B			Feature B			Feature B			Feature B		
Feature C	B1 (target)	B2	B3	Feature C	B1 (target)	B2	B3	Feature C	B1 (target)	B2	B3
C1 (target)	3	3		C1 (target)	3	3		C1 (target)			
C2	3		3	C2			3	C2		3	3
C3				C3		3	3	C3	3	3	3

Feature A											
Feature value A1 (target)				A2				A3			
Feature B			Feature B			Feature B			Feature B		
Feature C	B1 (target)	B2	B3	Feature C	B1 (target)	B2	B3	Feature C	B1 (target)	B2	B3
C1 (target)	3	3	3	C1 (target)	3	3		C1 (target)	3		
C2	3	3		C2	3		3	C2		3	3
C3	3			C3		3	3	C3		3	12

Results

Participants performed very well on the task, with an average of 0.1 % of misses (1 miss per 357 Go-trials per participant) and 1.3 % of false alarms (9.25 false alarms per 357 NoGo-trials per participant). They maintained a mean response time (RT) of 373 ms with a standard deviation (std) of 91 ms across conditions.

Response times varied greatly across conditions and participants. Participants mean response times varied between 400 ms (std: 90 ms) and 314 ms (std: 85 ms). Mean response times in conditions where only one target attribute was presented (colour only: c, form only: f, or size only: s) were 392, 382, and 485 ms respectively (std: 79, 78, and 107 ms respectively). In double-redundant conditions, i.e. conditions with two target attributes present (colour and form: cf, colour and size: cs, or form and size: fs), mean RTs were 319, 342, and 373 ms respectively (std: 46, 55, and 78 respectively). In the triple-redundant condition (all three target attributes present, cfs) the mean RT was 322 ms (std: 64 ms). Note that the variation between participant means is almost half as large as the difference in means of the slowest (s) and fastest (cfs) conditions.

Figure 2 a) shows the cumulative response time distributions of one representative participant, for all three single-target conditions, as well as for all three conditions where two target attributes were presented simultaneously, and the triple redundant condition. The probability of responding at time t or faster is plotted as a function of time. All participants were significantly faster in the

double redundant conditions cf and cs than in their constituent single target conditions (mean value of the Anderson-Darling test: cf vs. c: AD = 20.08, cf vs. f: AD = 16.69, cs vs. c: AD = 10.44, cs vs. s: AD = 26.53; the critical value of the AD test being 2.49 for a type I error rate of .05). All participants also responded significantly faster to form and size (fs) than to size alone (mean AD = 18.27). However, only participant 2 was significantly faster in condition fs than when only form was present as a target attribute (AD = 3.02). In the triple redundant condition, all participants were significantly faster than in the double redundant condition fs (mean AD = 12.44), three of four participants were faster than in condition cs (mean AD = 6.02), and none of the participants was faster than in condition cf (mean AD = 1.11).

Figure 2 b) shows the cumulative response time distributions for a representative participant of the three double redundant conditions as well as the triple redundant condition. Additionally it shows the Townsend Bound for each of these conditions. The Townsend Bound gives the upper limit of race model performance, based on the RT distributions of this participant in the single target conditions. All four participants respond significantly faster than the Townsend Bound in the triple redundant condition (mean AD = 6.73), as well as in the double redundant conditions cf (mean AD = 9.00) and cs (mean AD = 7.25). However, in the condition fs participants did not pass the Townsend Bound (mean AD = 1.17).

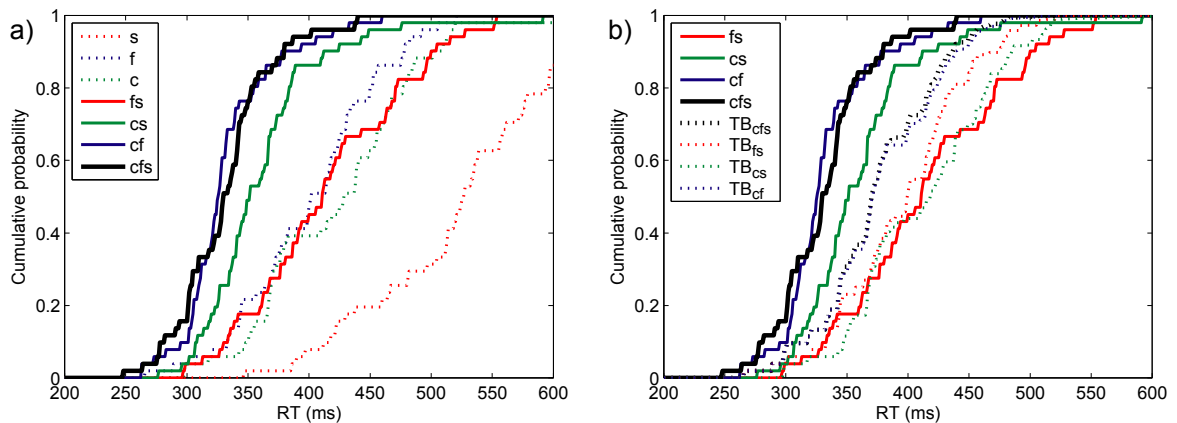
Figure 2. Experiment 3RedA: Performance of participant 3.

a) RT distributions for single, double and triple redundant conditions.

Dotted coloured lines are the cumulative distributions for single-target conditions, full coloured lines for double redundant conditions, black for the triple redundant condition.

b) Townsend Bounds for triple and double redundant conditions.

Coloured lines are double redundant RT distributions, the black line the triple redundant condition. Dotted lines of the same colour are the Townsend Bounds for the respective conditions.



As was to be expected from previous studies on redundancy gain, we managed to replicate findings of double redundancy gain that surpass the performance predicted by race models. However, we have not found evidence of a triple redundancy gain.

There are several patterns to be observed in these results. First, the attribute size is significantly slower than either colour or form for all participants. Second, participants perform equally well when both target attributes form and size are present than when colour only or form only are present. Third, performance on triple redundant trials and performance on trials with colour and form present does not differ significantly for any participant.

We therefore conclude that even though the Townsend Bound was violated for all participants in the triple redundant condition, these results cannot be evidence for a possible triple redundancy gain, nor for coactivation as an explanation of such a gain. Since the RT distributions for conditions cfs and cf do not differ in speed, all the gain in the triple redundant condition can be attributed to the presence of colour and form as target attributes. The presence of size does not seem to contribute to an additional gain.

In order to be able to observe a triple redundancy gain, it might be important that all target attributes have the opportunity to contribute equally to such a gain. If the recognition of one attribute is already much slower than the other two, this attribute can only contribute minimally, if at all, since the

advantage of its presence is negligible compared to the presence of two easily recognisable attributes.

This leads to the question why response times for size were slower than for colour or form. The most noticeable difference is that colour and form are absolute values, whereas size is a relative measure. The value of “large” or “small” only takes on meaning in relation to a reference, whereas “red” or “square” can be defined without a reference. In all further experiments we therefore chose only to investigate absolute target attributes.

We hypothesized that single target response time distributions need to be as close as possible to the same speed to be able to observe maximal redundancy gain. If one target attribute is processed noticeably slower than the other two, this attribute cannot contribute sufficiently to triple redundancy gain. There are a number of underlying processes involved in the processing of a visual stimulus (e.g. detection, identification, decision, and motor response). The speed of each of these processes, depending on the speed of the attribute to be processed, cannot be estimated, nor can their mutual independence be established. Therefore it is not possible to estimate *a priori* the processing speed of any given visual attribute. The most practicable solution to this problem is to test empirically a large number of different attributes until finding a set of three attributes which are processed at approximately the same speed. In this case, we would predict that triple redundancy gain is larger than double redundancy gain.

Further research: General method

The procedure and design of all further pilot experiments, designed to find an appropriate set of target attributes, stayed essentially the same as for the first experiment. Stimulus size was maintained at 3°VA (degree of visual angle). However, the stimulus attributes change for each experiment. Also, all further experiments use a different distribution of stimuli than the one used in 3RedA (and as a result of this, a different number of blocks per experiment and trials per block). Finally, presentation times of stimuli, as well as the time frame for participants to respond, were considerably shortened.

For all further experiments we used the stimulus distribution illustrated in Table I (bottom). This was done for two reasons. First, we felt it was necessary to include all possible combinations of attributes. In the distribution proposed by Mordkoff and Yantis (1991), target attributes were only combined with other target attributes or one of the two possible non-target values. The other non-target value was automatically associated with a non-target. In the case of the stimulus distribution used in this experiment, this meant that certain target attributes (such as the colour red) were never combined with one of the two possible non-target values of one of the other two stimulus dimensions (i.e. form or size). In order to reduce the impact of the identity of the non-target attribute, we felt that it was important for the non-target attribute(s) on single and double redundant targets not to be predictable. Secondly, by combining all possible attributes, we had the means of evaluating the impact of certain contingencies mentioned by Mordkoff

and Yantis (1991, 1993), thereby calculating the contribution of crosstalk to redundancy gain. There are now two types of stimuli per double redundant condition, one with each type of non-target attribute, therefore we can compare RTs in the double redundant conditions depending on non-target type (see chapter two for details on this).

However, this choice made finding a distribution of stimuli with three different attributes that satisfied the criteria formulated by Mordkoff and Yantis (1991) much harder. In the end, in order to keep 50% of targets, with more combinations of target and non-target attributes, and avoid facilitatory contingencies, we accepted certain inhibitory contingencies. But, as mentioned above, these contingencies would potentially slow down recognition of redundant targets, and also, we had the means to evaluate their impact. Therefore this choice of a stimulus distribution is conservative with respect to our goal of attributing redundancy gain to coactivation, and can be considered valid.

As a result of the new distribution, subsequent experiments consisted of 16 blocks per experimental session, with 60 trials per block, for a total of 960 trials. Presentation time of the fixation point preceding the stimulus was shortened to 494 ms, stimulus presentation itself was shortened to approximately 750 ms (this varied slightly between experiments, depending on the stimulus). The time limit for participants to respond to a stimulus was set at 750 ms after stimulus onset, and the feedback slide was presented for 753 ms only.

Practice effects

It is conceivable that practice effects (Newell & Rosenbloom, 1981) could have an effect on the observed redundancy gain. Either response times on single target trials might improve, thereby leaving less room for redundancy gain, or treatment on single target trials might stay the same while participants become experts on recognition of redundant trials, i.e. experts at recognising certain *combinations* of attributes. Since the task was fairly easy (low error rates in experiment 3RedA), we judged that practice effects if any would be visible after the completion of two sessions of approximately 45 minutes each. This hypothesis was tested in the following experiment.

3RedB : Colour, form and letter

Method. Three female undergrads from the Université de Montréal, with normal or corrected-to-normal vision, were compensated with 8\$ for their participation. For the reasons mentioned above, the target attribute size was replaced by the attribute letter. Stimuli were 3°VA in size, and varied in colour (red, green, or blue) and form (circle, square, or triangle). Each shape contained a cut out letter (either an H, a U or a B) of 2.7 degrees of visual angle (see Figure 1 b). Target attributes were red, circle, and the letter B. The letter B as a target was chosen because it shares attributes with both non-target letters. This experiment consisted of two experimental sessions, each with 960 trials, which were completed over two consecutive days by three subjects. This was done to test for potential effects of training that might influence the amount of redundancy gain

observed. All other design and procedure details remain the same as mentioned above.

Results. Error rates were as low as expected, with 0.07 % of misses (0.33 misses per 480 Go-trials) and 0.69 % of false alarms (3.33 per 480 NoGo-trials) in session one, and 0.07 % of misses (0.33 per 480 trials) and 1.11 % of false alarms (5.33 per 480 trials) in session two. Neither the number of misses ($t(4) = 0, p = 1$) nor the number of false alarms ($t(4) = -0.59, p = 0.59$) differed significantly between sessions.

An analysis of variance of mean response times by condition and session reveals that all participants respond significantly faster in session two than in session one ($F(1,28) = 12.22, p < 0.002$), and faster in the triple redundant condition than in the double or non-redundant conditions ($F(6,28) = 3.26, p < 0.015$). However, there is no interaction between condition and session ($F(6,28) = 0.2, p = 0.97$). We conclude that although practice does have an effect on response times, performance increases equally independent of the degree of redundancy. Therefore there is no effect of practice on the amount of redundancy gain between conditions. All subsequent experiments will contain only one experimental session of the above-mentioned length.

Analogous to other experiments, analysis of redundancy gains is done only on session one. As expected, participants mean response times varied a lot (participant 1: 393 ms (std 68); participant 2: 478 ms (std 155); participant 3: 362 ms (std 69)), as did response times between conditions. Mean response times in

conditions where only one target attribute was presented (colour only: c, form only: f, or letter only: l) were 417, 482, and 459 ms respectively (std: 98, 132, and 163 ms respectively). In double-redundant conditions, i.e. conditions with two target attributes present (colour and form: cf, colour and letter: cl, or form and letter: fl), mean RTs were 389, 383, and 423 ms respectively (std: 85, 93, and 122 ms respectively). In the triple-redundant condition (all three target attributes present, cfl) the mean RT was 358 ms (std: 84 ms).

Figure 3 a) shows the cumulative response time distributions of one representative participant, for all three single-target conditions, as well as for all three conditions where two target attributes were presented simultaneously, and the triple redundant condition. The probability of responding at time t or faster is plotted as a function of time. Only one participant showed a significant double redundancy gain for all three double redundant conditions over all single target conditions (value of the Anderson-Darling test between 28.42 and 6.96; the critical value of the AD test being 2.49 for a type I error rate of .05). For the other two participants, the combination of form and letter was not significantly faster than letter only, the faster of the two single target conditions (mean AD = 1.01). The combination of colour and letter was significantly faster than colour, the faster of the two single-target conditions, for one participant, but not for the other (AD = 2.60 and AD = 0.66 respectively). In the triple redundant condition, all participants were significantly faster than in the double redundant condition fl (mean AD = 14.74), and two out of three participants were faster than in condition cl (mean AD = 4.26) and in condition cf (mean AD = 8.16).

Figure 3. Experiment 3RedB: Performance of participant 3.

a) RT distributions for single, double and triple redundant conditions.

Dotted coloured lines are the cumulative distributions for single-target conditions, full coloured lines for double redundant conditions, black for the triple redundant condition.

b) Townsend Bounds for triple and double redundant conditions.

Coloured lines are double redundant RT distributions, the black line the triple redundant condition. Dotted lines of the same colour are the Townsend Bounds for the respective conditions.

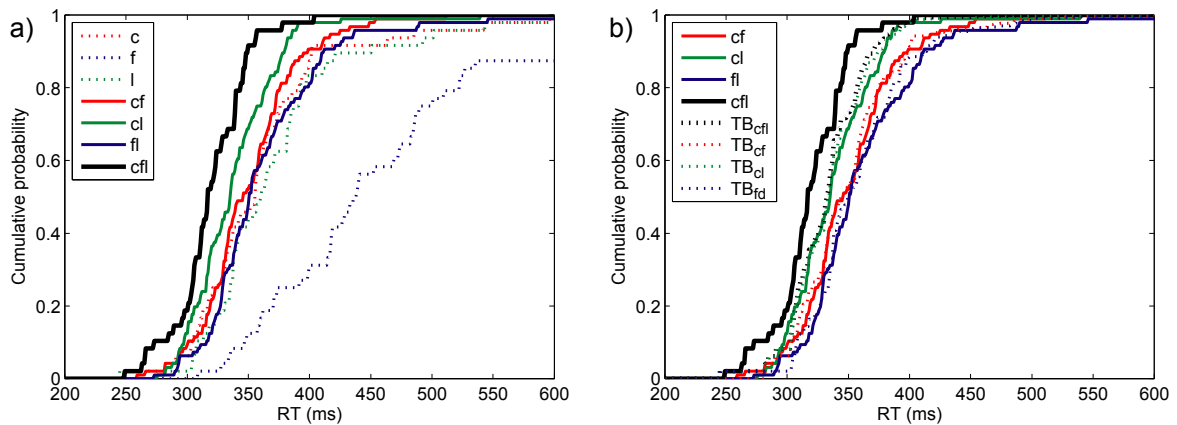


Figure 3 b) shows the cumulative response time distributions for a representative participant of the three double redundant conditions as well as the triple redundant condition. Additionally it shows the Townsend Bound for each of these conditions. The Townsend Bound gives the upper limit of race model performance, based on the RT distributions of this participant in the single target conditions. All three participants responded significantly faster than the Townsend Bound in the triple redundant condition (mean AD = 3.02). However, in all three double redundant conditions, only one participant was significantly faster than the corresponding Townsend Bound (cf: AD = 4.90; cl: AD = 6.62; fl: AD = 3.48).

Not all participants showed a significant redundancy gain in the double redundant conditions, let alone a gain large enough to overcome the Townsend Bound, and thus provide evidence against race models. Therefore we conclude that the gain observed in the triple redundant condition is again due to the interaction of two target attributes, without contribution of a third. This is supported by the fact that for all participants, the attribute form is processed slower than the other two attributes. Notably, when comparing the response time distributions of the double redundant conditions cf and fl to the distributions for single attributes, cf and c are practically overlapping (see Figure 3 a), as are fl and l, whereas form is visibly slower. Even at a double redundant level, one can conclude that form hardly contributes to an increase in performance. Since in experiment 3RedA form and colour were processed at approximately the same speed, we conclude that the relative decrease of performance for form is related to the addition of the attribute letter. This might be because letter and form are essentially two variations of the

same attribute, or because the attribute letter is more familiar (participants are exposed to letters innumerable times in everyday life), and might therefore be processed more easily (Wang, Cavanagh & Green, 1994).

Effects of masking

In an attempt to counteract the processing advantage the attribute letter has over form, stimuli were masked after a brief delay in the subsequent experiment. We hypothesised that since “reading” of a letter happens at a later stage in the processing pathway than identification of form (Kandel et al., 2000), a very brief exposure of the stimulus would force participants to process form and letter similarly, and thus equalise response time distributions of the two attributes. Another reason for masking was to avoid benefit from a visual imprint or afterimage left by the stimulus.

3RedC: Masked

Method. Six undergrads (3 male) from the Université de Montréal, with normal or corrected-to-normal vision, were compensated with 8\$ for their participation. Stimuli remained the same as for 3RedB (Figure 1b), except that the target colour was switched to green because this colour is less salient than red, in an attempt to render colour recognition slightly more difficult. The target form was also switched from circle to square, because a square shares attributes of each of the other two attributes, analogous to the target letter B, which shares characteristics of both non-target letters. Masks were constituted of four quadrants

from four randomly selected stimuli. Stimuli were presented for 75 ms, followed by a mask for 925 ms. All other methodological details are the same as mentioned above.

Results. Error rates were still low, although a little higher than in previous experiments, with 2.08 % of misses (10 misses per 480 Go-trials) and 4.03 % of false alarms (19.3 per 480 NoGo-trials). Mean response times varied between 364 ms and 519 ms across participants.

Mean response times over all participants in conditions where only one target attribute was presented (colour only: c, form only: f, or letter only: l) were 458, 476, and 485 ms respectively (std: 134, 149, and 125 ms respectively). Compared to experiment 3RedB, the difference between the mean RT of the fastest and slowest single-target condition has more than halved (3RedB: 65 ms; 3RedC: 27 ms). Processing speed of the three attributes seems to be more similar. The double-redundant conditions also had very similar mean RTs (cf: 409 ms (std 119 ms); cl: 420 ms (std 105 ms); fl: 419 ms (std 114 ms)). In the triple-redundant condition the mean RT was 384 ms (std: 100 ms).

Figure 4 a) shows the cumulative response time distributions of one representative participant, for all three single-target conditions, as well as for all three conditions where two target attributes were presented simultaneously, and the triple redundant condition. The probability of responding at time t or faster is plotted as a function of time. Five participants responded significantly faster in double redundant condition cf than in condition c (mean AD = 7.50; the critical

value of the AD test being 2.49 for a type I error rate of .05) and four participants faster than condition f (mean AD = 10.38). Two and four participants responded faster in condition cl than in conditions c (mean AD = 6.82) and l (mean AD = 11.27). Five participants responded significantly faster in condition fl than in f (mean AD = 5.04) or l (mean AD = 11.52). In the triple redundant condition, three, five and four participants responded significantly faster than in the conditions cf (mean AD = 5.10), cl (mean AD = 7.27), or fl (mean AD = 9.05) respectively.

Figure 4 b) shows the cumulative response time distributions for a representative participant of the three double redundant conditions as well as the triple redundant condition. Additionally it shows the Townsend Bound for each of these conditions. The Townsend Bound gives the upper limit of race model performance, based on the RT distributions of this participant in the single target conditions. Only one participant, in condition cf, performed significantly faster than the Townsend Bound for that condition (AD = 4.85). All other response time distributions in double or triple redundant conditions were not significantly different from their respective Townsend Bounds.

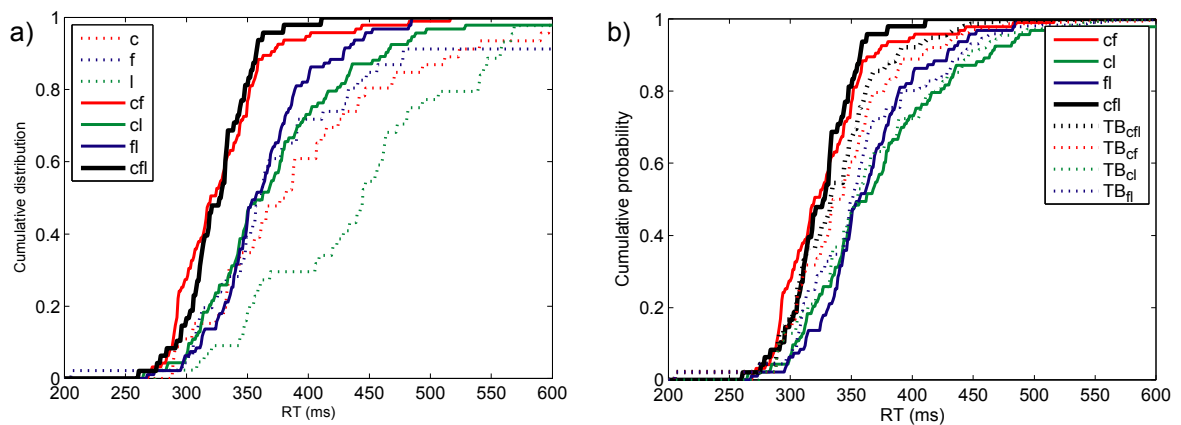
Figure 4. Experiment 3RedC: Performance of participant 3.

a) RT distributions for single, double and triple redundant conditions.

Dotted coloured lines are the cumulative distributions for single-target conditions, full coloured lines for double redundant conditions, black for the triple redundant condition.

b) Townsend Bounds for triple and double redundant conditions.

Coloured lines are double redundant RT distributions, the black line the triple redundant condition. Dotted lines of the same colour are the Townsend Bounds for the respective conditions.



The total absence of evidence in favour of coactivation models in this experiment is rather surprising, especially so since conditions for observing maximal redundancy gain were better here than in previous experiments (given our assumption is correct in that redundancy gain is maximal if single targets are processed as close as possible to the same speed). The variance between processing speed of single targets is smaller in this experiment than in previous ones, and processing speed for letter is similar to that of form – for some participants, form is even faster than letter.

We hypothesize that the disappearance of evidence against race models might be related to the masking of stimuli. To test for a causal relation between masking and the amount of redundancy gain observed, an experiment which balances masked and non-masked trials is conceived.

2RedM&M: Masked / Unmasked

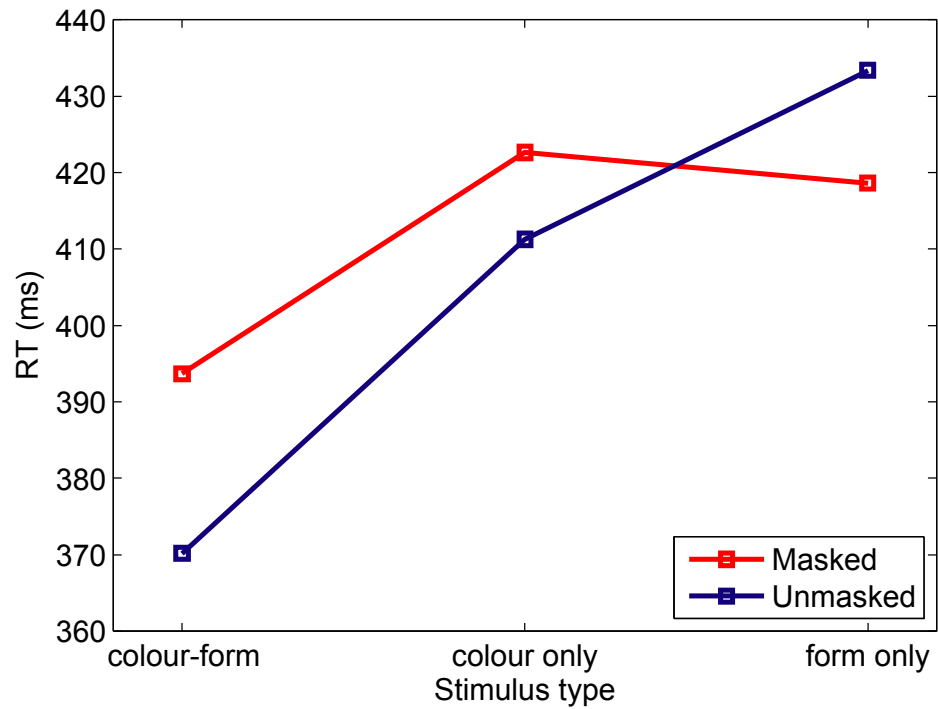
Method. To test the effect of a mask on redundancy gain, response times of four undergrads (1 female) from the Université de Montréal, with normal or corrected-to-normal vision, were measured in a Go-NoGo paradigm, with stimuli defined by *two* attributes, colour and form. Target stimuli were green and/ or square. Masks consisted of quadrants from four randomly selected stimuli. The experiment consisted of two experimental sessions, one in which stimuli were masked, the other in which they were unmasked, counterbalanced across participants. Sessions consisted of 576 trials each, half of which were non-target trials. Stimuli were presented for 1000 ms in the unmasked condition and for 75

ms, followed by a mask of 925 ms, in the masked condition. Analogous to the 3Red experiments, stimuli were preceded by a fixation point and followed by a feedback slide.

Results. An analysis of variance of mean response times by subject, stimulus type (redundant, colour only or form only) and session (masked or unmasked) reveals that although participants respond significantly faster to redundant than to non-redundant stimuli ($F(2,8) = 10.43$, $p < 0.006$), response times do not differ whether stimuli are unmasked or masked ($F(1,8) = 0.65$, $p = 0.44$), and there is no interaction between stimulus type and session ($F(2,8) = 1.84$, $p = 0.22$) (see Figure 5 for a plot of mean response times by session and stimulus type). We conclude that masking does not have an effect on response times in general, or on the amount of redundancy gain between conditions.

Figure 5. 2RedM&M: effect of condition and masking on response times.

Conditions are plotted on the x-axis. Mean response times when stimuli were masked are plotted in red, unmasked in blue.



In order to find an answer to two major problems in the previous experiments, we decided to look at the structure of the visual processing pathway in the brain. First, this might help us find an appropriate combination of three target attributes. Up to now, one attribute was always considerably slower than the other two, and therefore affected the triple redundancy gain. Second, we needed to explain the mixture of our results: for some combinations (pairs or set of three) of attributes we find redundancy gain significantly faster than the Townsend Bound, for others no redundancy gain, or gain but not above the Townsend Bound. This could also be related to the structure of visual processing areas.

Note that we do not wish to establish a causal relation between the structure of the visual system and our results. We merely use it as an indication to point us in the right direction. The main foundation for the conception of further experiments remains empirical data.

Early processing stages

Visual processing is organised in a hierarchical fashion, going from the analysis of single neuron responses to more and more complex units of information (Maunsell & Newsome, 1987). The primary visual processing area V1 is the first area of visual processing to receive input from both eyes, both from the Magnocellular and Parvocellular pathway (Hubel & Wiesel, 1979; Hubel, 1988). V1 is selective to orientation and spatial frequency information, which is used to define contours, one of the most basic features to be extracted from visual input (Hubel & Wiesel, 1962, 1979). Contours are what is essentially used in form

recognition, which means that form recognition can happen at a processing stage as early as V1. V1 is also selective to colour (Hubel, 1988; Livingstone & Hubel, 1987).

Colour and form seem to be the most reliable attributes across experiments (they are processed at roughly the same speed and produce redundancy gain fairly reliably faster than the Townsend Bound). If we break down form into its two components, spatial frequency and orientation, we get a set of three attributes which are all processed at a very early level in visual processing, and all in the same area, namely in V1. Since we know that colour and form are processed at roughly the same speed, we hypothesize that these three attributes will be as well.

3RedD : colour, orientation, frequency

Method. To test this hypothesis, four undergrads (1 male) from the Université de Montréal, with normal or corrected-to-normal vision, participated in an experiment similar to those described above, with stimuli being defined by three attributes: colour, spatial frequency of bars and orientation of bars. Stimuli consisted of squares (size: 3°VA) filled with a sinusoidal gratings alternating between colour and gray of different cycle length and orientation. Luminance of stimuli was reduced to 20%, and saturation to 50%. Stimuli were presented in front of a gray background with luminance reduced to 20%. This was done to slow down the recognition of target attributes, in particular colour, thereby leaving more room for improvement due to redundancy. Also, the attribute spatial frequency is again a relative instead of an absolute attribute. Therefore the disadvantage to

colour is also an effort to compensate roughly for the disadvantage of spatial frequency. The target colour was switched to blue because in low luminance and saturation, blue is less salient than red or green. Again, this is an effort to render colour recognition slightly more difficult. Target spatial frequency was a cycle of 0.25°VA (non-targets: 0.5 and 0.75°VA), and target orientation was diagonal (non-targets: vertical and horizontal). See Figure 1 c) for an example of stimuli. Stimuli were presented for 750 ms. All other methodological details are the same as mentioned above.

Results. Error rates were low, with 1.30 % of misses (6.25 misses per 480 Go-trials) and 1.20 % of false alarms (5.75 per 480 NoGo-trials). Mean response times varied between 407 ms and 467 ms across participants. The reduced luminance and saturation do not seem to have affected performance relative to previous experiments.

Mean response times in conditions where only one target attribute was presented (colour only: c, frequency only: f, or orientation only: o) were 432, 469, and 487 ms respectively (std: 82, 73, and 96 ms respectively, and did not differ significantly ($F(2,6) = 2.85$, $p = 0.11$). The double-redundant conditions also had very similar mean RTs (cf: 405 ms (std 54 ms); co: 409 ms (std 69 ms); fo: 434 ms (std 67 ms); $F(2,6) = 1.37$, $p = 0.30$). In the triple-redundant condition the mean RT was 394 ms (std: 52 ms).

Figure 6 a) shows the cumulative response time distributions of one representative participant, for all three single-target conditions, as well as for all

three conditions where two target attributes were presented simultaneously, and the triple redundant condition. The probability of responding at time t or faster is plotted as a function of time. All participants showed a significant double redundancy gain for all three double redundant conditions, with the exception of two participants, who were not significantly faster than for colour alone in the corresponding redundant conditions cf and co. All participants were also significantly faster in the triple redundant condition than in the condition fo (mean AD = 10.02; the critical value of the AD test being 2.49 for a type I error rate of .05). However, cfo and cf did not differ for any participant (mean AD = 1.61), and cfo and co differed for only one participant (AD = 2.90)

Figure 6 b) shows the cumulative response time distributions for a representative participant of the three double redundant conditions as well as the triple redundant condition. Additionally it shows the Townsend Bound for each of these conditions. Only in conditions cf and cd, one participant performed significantly faster than the Townsend Bound (AD = 2.75 and 2.52 respectively). All other response time distributions in double or triple redundant conditions were not significantly different from their respective Townsend Bounds.

Although the three single target attributes are processed at roughly the same speed, and although we are able to refute race models as an explanation for double redundancy gain, there is still no evidence for triple redundancy gain in this experiment. We therefore decided to try and increase the task difficulty again by reducing saturation to a minimum.

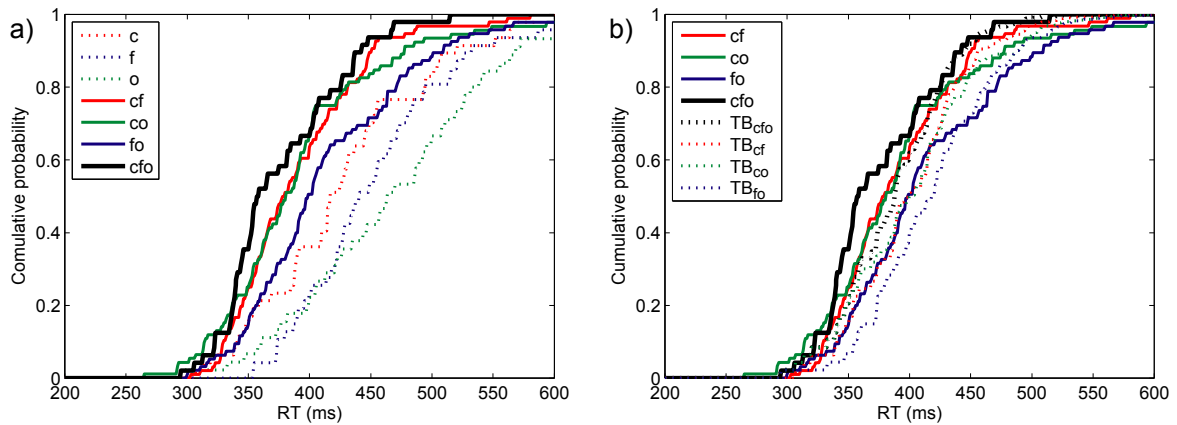
Figure 6. Experiment 3RedD: Performance of participant 2.

a) RT distributions for single, double and triple redundant conditions.

Dotted coloured lines are the cumulative distributions for single-target conditions, full coloured lines for double redundant conditions, black for the triple redundant condition.

b) Townsend Bounds for triple and double redundant conditions.

Coloured lines are double redundant RT distributions, the black line the triple redundant condition. Dotted lines of the same colour are the Townsend Bounds for the respective conditions.



3RedDSat: low saturation

Method. Four female undergrads from the Université de Montréal, with normal or corrected-to-normal vision, participated. Luminance of stimuli and background remained at 20%, whereas saturation of stimuli was reduced to 10%. All other details remained the same as in 3RedD.

Results. Although the percentage of misses stayed in the same range as for previous experiments (1.40 %; 6.75 misses per 480 Go-trials), the number of false alarms increased (3.70 %; 17.75 per 480 NoGo-trials). Apart from that, all result patterns from 3RedD were essentially reproduced. Mean response times between single target conditions did not differ ($F(2,6) = 2.19$, $p = 0.19$). With four exceptions, response times of all participants in double redundant conditions were significantly faster than in all single target conditions, and faster for cfo than for fo. Only one participant managed to respond faster to cfo than to both cf and co. Performance on double and triple redundant trials did not differ significantly from the Townsend Bound ($AD < 2.42$).

In light of the results of the two previous experiments, we decided to revert back to colour and form as two target attributes. We decided to add direction of movement as a third attribute. Motion is another feature which is extracted very early on from visual input, also as early as in V1, and mainly in the medial temporal cortex MT (Maunsell & van Essen, 1983; Born & Bradley, 2005). Visual input is divided depending on type, and processed along two main processing pathways: Motion is processed in the dorsal “Where” pathway, whereas colour and

form are processed mainly in the ventral “What” pathway (Goodale & Milner, 1992), and both are largely independent of each other (Livingstone & Hubel, 1987).

3RedE1: colour, form, and movement

Method. Eight participants took part in this experiment. Stimuli are defined by three features: colour, form and direction of movement, target attributes being blue, square, and/ or movement to the right. Luminance of stimuli was reduced to 20%, and saturation to 30%. Stimuli were presented in front of an equiluminant gray background. Stimuli started out at the center of the screen, and then moved outward for 4.5°VA , at an angle of 45° (target), 165° or 285° . Target stimuli were blue and/ or square and/ or moved to the right. See Figure 1 d) for an example of stimuli, and chapter two for further details on methods.

Results. There is no significant difference between mean response times for targets defined by colour, form or direction of movement ($F(2,14) = 0.54$, $p = 0.59$). The majority of subjects showed significant triple and double redundancy gain, often significantly faster than predicted by the respective Townsend Bound. In order to get more data, an additional 16 participants were measured in the same experimental setup. Since there were no significant differences between error rates and mean response times of all 24 participants (see chapter 2), they will be analysed together. For more details on method and results, refer to chapter two.

Do certain brain areas favour coactivation?

Another hypothesis we pursued for a while was that coactivation is specific to certain brain regions, and therefore we found evidence of coactivation for some attributes, but not for others. There is evidence that the visual area V4 for instance responds selectively to combinations of colour and form, or that response to form is modulated by information on colour (Zeki, 1977; Motter, 1994). This could be a possible explanation why the combination of colour and form worked better than other attribute combinations to produce a redundancy gain. Other areas in the visual cortex where evidence exists that they respond selectively to combinations of features include the inferotemporal (IT) cortex (colour and shape; Lueschow, Miller & Desimone, 1994; Komatsu & Ideura, 1993), MT (speed and direction; Maunsell & van Essen, 1983), and the medial superior temporal cortex MST (perceived depth and direction of motion; Roy, Komatsu & Wurtz, 1992; Graziano, Anderson & Snowden, 1994).

It is not surprising that later areas on the visual processing pathway such as the IT or MST are selective to combinations of features. The further along an area is situated on the processing pathway, the more complex the features it is selective to become, i.e. they are constituted of several less complex features as their individual components. Area MST is selective to the optical flow of objects, i.e. the apparent motion of objects when the observer moves (Duffy & Wurtz, 1995). The apparent motion depends on the degree of binocular disparity or perceived depth and the perceived direction of movement.

In order to test whether coactivation could be correlated to combined selectivity of features, we conceived a target detection task with two target attributes other than colour and form to which a combined selectivity exists. Should we find evidence for such a correlation, we predict that triple redundancy gain is only possible for features that a specific brain area is selective to in combination.

2RedMST : perceived depth and direction of movement

Method. To study the redundancy gain due to perceived depth and direction of movement, response times of eight undergrads (5 male) from the Université de Montréal, with normal or corrected-to-normal vision, were measured in a Go-NoGo paradigm, with stimuli defined by *two* attributes, perceived depth and direction of movement. Participants were selected due to their ability to perceive depth using red / cyan 3D glasses, and were required to wear these glasses throughout the experiment.

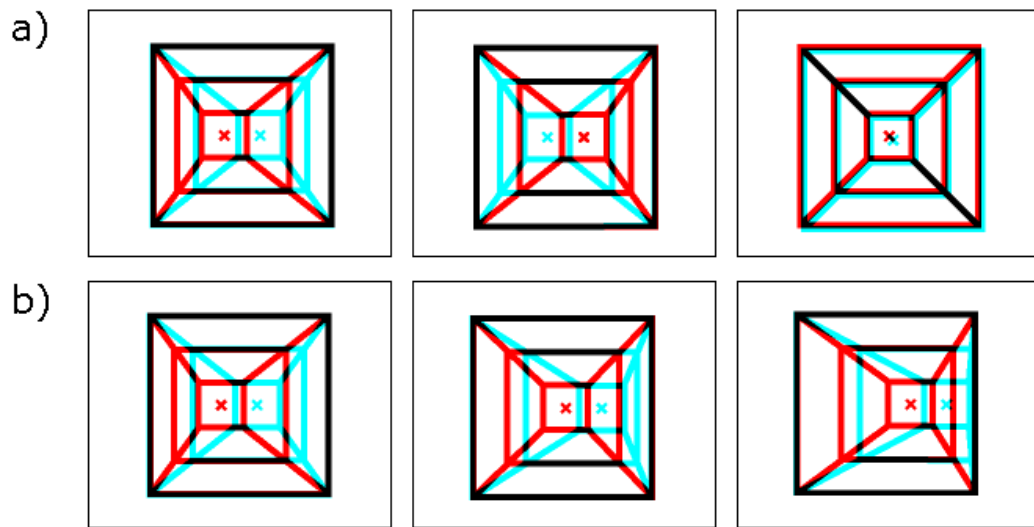
Stimuli measured 3°VA, and consisted of two superposed images of two squares (one inside the other and connected at the corners), one in red and one in cyan. They were conceived such that while wearing red / cyan 3D glasses, the outer square appeared level with the computer screen and the inner square seemed to be either in front of, level with or behind the computer screen (see Figure 7 a) for examples of the different degrees of perceived depth of stimuli). Participants were told stimuli resembled either a pyramid (inner square in front) or a hallway (inner square behind screen) or were flat. Stimuli moved outwards from the center

of the screen for 6°VA during ten frames at an angle of 90° , 180° , or 270° . The relative position of the inner square with respect to the outer square changed progressively with the movement, in order to maintain the illusion of three-dimensionality. See Figure 7 b) for an example of a pyramid moving to the right (first, middle, and last frame). Target stimuli were pyramids and/ or moving to the left.

The experiment consisted of eleven blocks with 63 trials per block. Response times and performance for the first 15 trials per block (nine Go trials) was not measured, as these were considered training trials. 50% of the 48 test trials were Go trials, 8 of which were redundant target trials. For each trial, participants viewed a blank screen for 490 ms (followed by a text message for 644 ms in the training trials, telling participants if the next stimulus was a pyramid, a hallway, or neutral), a fixation point level with the screen for 490 ms, the stimulus at the screen center for 294 ms, followed by ten frames of outward movement, the first nine for a duration of 42 ms, and the last frame lasting 98 ms, and finally a feedback slide for 742 ms. Participants were required to respond to Go trials faster than 775 ms. All other details were similar to previous experiments.

Figure 7. Examples of stimuli for the experiment 2RedMST

- a) **Three attributes of the feature depth**, from left to right: in front of screen (target), behind screen, and on the same level as screen
- b) **Example of optical flow for a stimulus in front of screen moving to the right**. From left to right: initial frame, presented in screen center; middle frame, presented at 3° VA of screen center; last frame, presented at 6° VA from center of the screen.



Results. Error rates were very high, with 37.14 % of misses (140.38 misses per 396 Go-trials) and 15.18 % of false alarms (57.38 per 360 NoGo-trials). Mean response times varied between 437 ms and 614 ms across participants. Both high error rates and high between-participant variability in mean response times reflect the difficulty of the task. Figure 8 a) shows mean response times per subject and condition. Mean response times when only the target attribute depth was present (mean: 479 ms, std: 140) were significantly faster than when only movement was present (mean: 592 ms, std: 75 ms; $t(14) = -2.69$, $p < 0.01$). All participants showed this pattern, except one, for whom movement recognition was faster than depth recognition. In the redundant condition the mean RT was 482 ms (std: 114 ms). Figure 8 b) shows the cumulative distributions of a representative participant in the single target and redundant conditions. All but one participant performed significantly faster in the redundant condition than for movement only (mean AD = 19.91), and only this participant performed significantly faster than for depth only (AD = 32.23), all others did not (mean AD = 1.35). None of the participants violated the Townsend Bound in the redundant condition.

All participants seem to rely on one feature (depth for seven out of eight participants) to recognise target objects. Only when this has been identified as a non-target do they move on to the second feature. This explains why RT distributions for the “dominant” attribute and the redundant condition are virtually superimposed for all participants, and why no redundancy gain is observed. This pattern is possibly caused by the increased difficulty of the task, which seems to force a serial instead of parallel processing of attributes, thus making redundancy

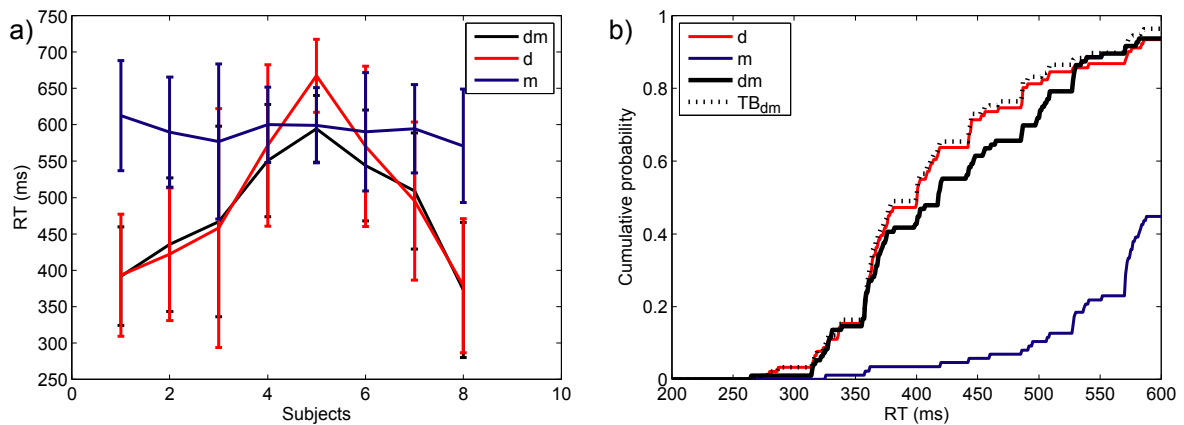
gain impossible (Van der Heijden et al., 1983, Krummenacher, Müller & Heller, 2001). In view of these results, we cannot confirm or reject the hypothesis of combined selectivity to explain coactivation. We therefore turn from trying to find a biological explanation of our results to modelisation. Additionally, in order to gain more information on which to build possible models of redundancy gain, we investigate the characteristics of redundancy gain more closely.

Figure 8. Experiment 2RedMST.

a) Mean response times per subject and condition. Error bars represent std.

Participants are plotted on the x-axis, the red line plots the mean RTs for depth recognition, the blue line for direction of movement, and the black line for redundant targets

b) RT distributions and Townsend Bound for single and double redundant conditions (participant 2). The red line is the cumulative RT distribution for depth, blue for direction of movement, black for redundant targets. The black dotted line is the Townsend Bound.



Characteristics of redundancy gain

The universally acknowledged effect of redundancy gain is a decrease of response times to a redundant target as opposed to non-redundant targets. However, there could well be other manifestations of redundancy on response times, which are not related to a decrease. We investigated this possibility using the data from experiment 3RedE, the main experiment of this thesis, with target attributes colour, form and direction of movement (see chapter 2 for details, and above for details on the corresponding pilot, 3RedE1). We fitted Weibull distributions to response time distributions for all participants and conditions, and analysed parameter changes between redundant and non-redundant conditions. Weibull distributions are characterised by three parameters: 1) shift, the onset, or minimum of the distribution; 2) scale, the range of the distribution; and 3) symmetry, the degree to which tail ends of a distribution are symmetrical. Redundancy gain could manifest itself in a change of any or all of these characteristics.

When comparing parameters over conditions, we found that shift decreases for redundant targets, scale remains constant, and symmetry increases. Race models predict shift to remain constant. They are not compatible with minimal response times decreasing in redundant conditions, and therefore are not compatible with the parameter variations observed here. However, the predictions of coactivation models are less clear; therefore we tested them in a series of Monte Carlo simulations. These simulations showed that coactivation models predict a

decrease in shift. However, they cannot account for an increase in symmetry, and are therefore not compatible with the pattern of parameter change observed in the experimental data either. Further details on fitting and simulation methods and results can be found in chapter two.

In conclusion, redundancy gain is characterised by a decrease in mean response times, a decrease in minimal response times in particular, and an increase in symmetry of the distribution of response times. This pattern is not fully predicted by either race models or coactivation models. Therefore, the possibility of a model that does explain this pattern needs to be investigated.

Modelling

Studying the structure of the visual processing pathways did not provide insights as to why we observed a violation of the Townsend Bound in some cases but not in others. We initially turned to modelling in order to find an alternative, unifying model, which fitted all our results. The experimental data from pilots (see above) and main experiment (see chapter two) does not point conclusively to a single causal explanation for redundancy gain, be it race models, crosstalk or coactivation. Rather, the results seem to favour one theory for one set of target features, and another for a different combination of features. We therefore decided to construct a single alternative model to account for all the data. A parsimonious, unified approach would be preferable to multiple explanations, as target features are presented unimodally. If a new model, in addition to already existing ones, is to make sense, we would expect it to predict absence of redundancy gain in some cases, *and* violation of the Townsend Bound in others, as well as the decrease in shift and increase in symmetry of response times in redundant conditions, thereby also providing a unified explanation for often contradictory literature on redundancy gain.

First attempts at modelling our data were based on exemplar-based memory models (Estes, 1986; Nosofsky, Clark & Shin, 1989; Cohen & Nosofsky, 2000). By modifying the level of noise with which different features of an exemplar stimulus are encoded, we tried to modulate the recognition rate, thus

simulating an RTE. These attempts were not successful, and we did not continue efforts in this direction.

We developed the cascade race model, a model where channels are triggered by a cascade of activations before satisfying a single decision criterion (hence the name cascade), to explain our results. Indeed, the cascade race model is consistent with an absence of redundancy gain as well as a violation of the Townsend Bound; it predicts an increase in symmetry and a decrease in minimal response times for redundant conditions (further details on parameters and predications of the cascade race model in chapter two). We therefore propose the cascade race model as a unifying explanation of response time behaviour for redundant targets.

Final research questions

During the series of pilot experiments described above, the following research questions emerged. These questions will be explored in the subsequent two chapters, and in their answer lies the main contribution of this thesis:

- (1) Is triple redundancy gain possible? If so, does it exceed the Townsend Bound (in particular the Townsend Bound for mixed models)?
- (2) Does redundancy gain manifest itself in other ways than a shift in mean of response time distributions (i.e. change in symmetry, scale, or at the tail ends of a distribution)?
- (3) Which statistical test is the most powerful and accurate for detecting changes in response time distributions?
- (4) Can coactivation models explain all observed differences in response time distributions?
- (5) If not, can we find a model which does explain the entirety of our results?

CHAPTER 2

Coactivation results cannot be explained by pure coactivation models

Sonja Engmann and Denis Cousineau

Université de Montréal

Manuscript submitted to the journal *Attention, Perception and Psychophysics*

Déclaration des coauteurs d'un article

1. Identification de l'étudiante :

Sonja Engmann

Ph. D Psychologie, option Sciences Cognitives et Neuropsychologie

2. Description de l'article :

Engmann, S., & Cousineau, D. (submitted). **Coactivation results cannot be explained by pure coactivation models.** Submitted to *Attention, Perception & Psychophysics*.

3. Déclaration de tous les coauteurs autres que l'étudiante :

À titre de coauteur de l'article identifié ci-dessus, je suis d'accord pour que Sonja Engmann incluse l'article identifié ci-dessus dans sa thèse de doctorat, qui a pour titre : « Redundancy gain : manifestations, causes and predictions ».

Denis Cousineau
Coauteur

Signature

Date

Apport original

Apport original de Sonja Engmann à l'article :

“Coactivation results cannot be explained by pure coactivation models.”

J'ai préparé et effectué les expériences présentées dans cet article, ainsi que les analyses, sous la supervision et avec les conseils de Denis Cousineau. L'écriture de l'article a été faite en collaboration avec le Dr Cousineau.

Abstract

Response times in a visual object recognition task decrease significantly if targets can be distinguished by two redundant attributes. Redundancy gain for two attributes has been commonly found, but redundancy gain from three attributes has been found only for stimuli from three different modalities (tactile, auditory, and visual). This study extends those results by showing that redundancy gain from three attributes within the same visual modality (color, form and direction of movement) is possible. It also presents evidence that neither race models nor coactivation models can account for such a gain, and introduces a novel method of detecting triple redundancy gain that surpasses race model predictions. Finally, the cascade race model is introduced. This new model can explain the results from this study as well as previous research on redundancy gain. It thus provides a unifying account of the redundant target effects.

Introduction

The environment surrounding us can be subdivided into distinct sources of information that are used to make a decision about the identity of objects. Sources of information are from different modalities (e.g. auditory, visual) with different types of information within each modality (e.g. color, form, direction of motion within the visual modality). These features are processed and identified separately (although not necessarily independently) by the visual system through specialized processing channels before the object is perceived as a whole (Treisman & Souther, 1985; Kandel, Schwartz & Jessell, 2000; Ungerleider & Mishkin, 1982; Milner & Goodale, 1993).

In some cases, a single feature (e.g., the color) is sufficient to recognize an object. Treisman & Souther (1985) have shown that if a target object differs from several distracters by one distinct feature alone (e.g., a red square among green squares) it can be detected rapidly, accurately and without conscious effort. The detection is also independent of the number of surrounding distracters. This is known as the pop-out effect (Treisman & Souther, 1985).

In other cases, a combination of several features is needed for an unambiguous identification. If the joint identification of two or more features is necessary to distinguish a target object from several distracters (e.g., a red square among green squares and red circles), target recognition becomes slower and error-prone. Even if the target itself is unique among the distracters, it shares at

least one feature with any one of the distracters, which makes it less easily distinguishable. This task requires central attention and its difficulty increases proportionally with the number of distracters.

In yet other cases, target detection or recognition is facilitated by the presence of multiple target attributes. If a target object is defined by several features and the presence of either one of them on its own – as opposed to a combination of all target features – is sufficient to unambiguously recognize the target, target recognition is faster when more than one target feature is present. For example, red squares (i.e. targets with both features) will be detected faster than red circles or green squares (i.e. targets with only one of the two features). This is known as the Redundant Signals Effect (Kinchla, 1974) or the Redundant Target Effect (RTE; Miller, 1982).

The Redundant Target Effect (RTE) is a phenomenon that has proven to be consistent and stable whenever attention needs to be divided among several modalities, locations or feature dimensions and when several input channels separately provide the necessary information to perform a task (Miller, 1982; Van der Heijden, La Heij & Boer, 1983; Kinchla & Collyer, 1974; Van der Heijden, 1975). Bimodal and even trimodal detection tasks show facilitation if a stimulus is presented on several different modalities more or less simultaneously (Bimodal: Wundt, 1880; Fidell, 1970; Mulligan & Shaw, 1980; Miller, 1982; Trimodal: Van der Heijden et al., 1983; Diederich & Colonius, 2004, Krummenacher, Müller & Heller, 2001, Miller, 1981, Marzi et al., 1996).

The Redundant Target Effect generalizes across target dimensions – form, color, orientation, etc. (Miller, 1982; Mordkoff & Yantis, 1993; Feintuch & Cohen, 2002), as well as letters and words (Morton, 1969) – and modalities (visual, auditory, and tactile; Diederich, 1995). The fact that reaction times profit from redundant signals, at least under most conditions and tasks, provides sound evidence that parallel rather than serial processing of input does happen. The RTE cannot be explained without assuming parallel processing at some stage of the processing pathway (Van der Heijden et al., 1983; Krummenacher et al., 2001).

However, several factors influence the size or the appearance of the RTE. Some of these are linked to the experimental design – for example distracter presence or absence. If only one stimulus is present during single target trials, then the RTE is much smaller than if a distracter is present on the other channel on single target trials (Miller, 1982; van der Heijden, Schreuder, Maris & Neerinx, 1984). It seems that attention focused on one channel – as opposed to divided attention in cases where two or more channels have to be monitored – is sufficient to reduce, or in some cases, to completely compensate for any redundancy gain (Miller, 1982). The type of task is also important: redundancy gain is typically observed in experimental paradigms of the type Go-NoGo, where a response is required of the participant if and only if any one of several redundant features is present (Miller, 1982; Mordkoff & Yantis, 1991, 1993; Diederich 1995). Redundancy gain has also been found in a two-alternative-forced-choice paradigm (2AFC; Fidell, 1970), but it is not as large as in Go-NoGo paradigms (Grice & Reed, 1992). Also, it is not entirely clear whether a gain in 2AFC paradigms is

really due to parallel processing of redundant target features (van der Heijden et al., 1983).

The spatial location of two redundant visual targets affects RTE as well: the farther they are apart, the lesser is the redundancy gain (Feintuch & Cohen, 2002; Colonius & Diederich, 2004). However, if targets in spatially different locations are bound together by grouping, the redundancy gain can be increased considerably, as they are then perceived as belonging to the same object (Feintuch & Cohen, 2002).

Several types of models have been proposed to explain the redundant target effect: race models, coactivation models and various hybrids of these two. Both race and coactivation models are generally based on the assumption of independent channels that contribute to the accumulation of evidence. It is rather unlikely, however, that this assumption holds in reality. Mordkoff and Yantis (1991) showed that activity on one channel can be influenced by events on another channel, and several authors introduce lateral inhibition between channels to be able to explain their results on various reaction time tasks (Usher & McClelland, 2001; Huber & Cousineau, 2004).

The race model was one of the first models proposed to explain the RTE (Raab, 1962). It assumes independent channels separately accumulating evidence in favor of the specific signal or feature to which they are tuned. As soon as one of the channels has accumulated enough evidence to surpass the decision threshold, this channel – the fastest, hence the name of the model – determines the output of

the model. The RTE is explained by the notion of statistical facilitation first introduced by Raab (1962). The author showed that when sampling random reaction times across channels, the distribution of the minimal response time of each of these samples will always have a mean lower than any of the response time distributions of the different channels. However, the starting point of the minimal response time distribution cannot be lower than the minimal response time across channels, but the variance of the minimal RT distribution will be smaller than the variance of any of the individual channel distributions.

Coactivation models were developed as an alternative to race models for explaining the RTE (Smith, 1968; Miller, 1982; Schwarz, 1989). Coactivation is defined as an activation build-up from different channels to satisfy a *single* threshold criterion. Coactivation models differ chiefly from race models in that the activation from the different channels is combined at some point in the processing of the input. Activation from all channels jointly determines what the response at the next processing level should be. In fact, it is the joint activation of a single threshold criterion from different channels which enables coactivation models to predict a redundancy gain: Even if activation on any one channel alone is insufficient to overcome the threshold and make a decision, the pooling of activation from several still weakly activated channels makes it possible to overcome the threshold faster than with any single channel alone.

Various authors have compared separate activation and coactivation models (e.g., Mulligan & Shaw, 1980; Fidell, 1970; Kinchla & Collyer, 1974;

Eriksen & Schultz, 1979), but the conclusions are not homogeneous and are often contradictory (Mulligan and Shaw, 1980; Fidell, 1970; Eriksen & Schulz, 1979).

Several attempts have been made to find a criterion that allows a conclusive distinction between race models and coactivation models. A possible way of excluding separate activation models irrefutably was proposed by Miller (1978). The performance of race models on redundant target trials is simply the minimum response time of the different channels that contribute to the redundant signal (Raab, 1962). This allows us to calculate the best possible performance of race models. Two of the most well-known methods to calculate the upper limit of race models are the Miller Inequality (Miller, 1978) and the Townsend Bound (Townsend and Nozawa, 1995). If response time distributions on a redundant target task exceed either of these criteria, race models can be refuted as an explanation for redundancy gain; they are not capable of accounting for the amount of gain induced in redundant target trials.

The Miller Inequality has been used frequently to refute race models as the sole explanation for the redundancy gain in detection tasks with targets from different dimensions (Krummenacher, Müller & Heller, 2002), different modalities (Diederich & Colonius, 1987), and letter search tasks (Miller, 1982). Even in participants with lateral visual extinction, the RTE induced by a stimulus in the extinct hemisphere was strong enough to violate the Miller Inequality (Marzi et al., 1996). In extending the Miller Inequality to include three redundant targets, Diederich and Colonius (2004) found evidence to refute race models in trimodal

detection tasks: the gain observed between double redundant and triple redundant targets alone was too large to be explained solely by separate activation models.

Violation of the Miller Inequality, the Townsend Bound (see results) or any other criterion defining the upper limit of race model performance (e.g. Grice, Canham & Gwynn, 1984) is usually interpreted as evidence of coactivation somewhere along the processing pathway. However, Mordkoff and Yantis (1991) suggested an alternative explanation: crosstalk. With the Interactive Race Model, they proposed an extension of separate activation models, which integrates inter-channel crosstalk (positive or negative contingencies between target and target, target and non-target, or two non-targets on two channels) and bias towards one response.

Mordkoff and Yantis (1991) explain Miller's (1982) and similar results in terms of existing contingencies between stimuli on different channels. In a series of experiments with a letter search task they then show that the Miller Inequality is not violated if all contingencies between channels are equated.

In a rigorous test of the interactive race model (Schwarz, 1996), most of the results from Mordkoff and Yantis (1991) were replicated. However, under certain conditions (non-simultaneous signal presentation) violation of the Miller Inequality was consistently found even when inter-channel contingencies were equated. Miller (1981) also obtained violation of the Inequality in the absence of inter-channel correlations, as did Mordkoff and Yantis (1993). The latter concluded that although inter-channel crosstalk does influence response times,

coactivation is at least partly responsible for facilitation in cross-dimensional redundant targets, whereas within the same dimension, separate activation is sufficient to explain facilitation.

In conclusion, the literature is less than clear or unanimous as to possible causes of redundancy gain. It is therefore important to find a unified approach to causes of redundancy gain, something we try to achieve with this study.

The present study pursues two different goals. First, we wish to investigate if redundancy gain from three redundant target attributes inside a single modality is possible. To the best of our knowledge, triple redundancy within any single modality has never been addressed before. One reason for expanding the study of redundancy gain in that direction is that the ecological validity of the target paradigm increases. In a natural context, we rarely see targets which are defined by only two target attributes. Therefore, increasing the cognitive load and studying triple redundancy gain is likely to reveal interesting insights into the processing of visual stimuli. Another reason for choosing a triple redundant paradigm is that it will likely give valuable information about the dynamics of a redundancy gain. It might give us an idea about an upper limit to gain in RTs, limits with respect to the type of target attributes which can induce triple redundancy gain, and the factors that could hide or inhibit redundancy gain.

The second goal of this study is to differentiate between possible causes of redundancy gain. As mentioned above, the literature is not at all unified in attributing redundancy gain to statistical facilitation, to crosstalk or to coactivation.

Based on our experimental data, we will exclude all three of these as an isolated explanation of the RTE. Since crosstalk is likely to exist, we exclude it by allowing no possible facilitatory contingencies between target attributes on different channels. In a second step, we will reject statistical facilitation by showing that performance of participants on redundant target recognition will be significantly above the Townsend Bound. In a final step, we will also refute coactivation models as an explanation for redundancy gain by comparing the minimal response time as well as standard deviation and skew of participants' response time distribution with the coactivation model's predictions for those response time distribution characteristics. .

We will propose a cascade race model of statistical facilitation as an alternative explanation of redundancy gain. We will show that this model is able to explain all the characteristics of our experimental data, and further provides a unifying and unambiguous explanation for redundancy effects in literature.

Method

Participants

Participants were 24 undergraduate students (17 females) from the Université de Montréal, between 19 and 27 years of age. All had normal or corrected-to-normal vision. Participants were compensated with 8\$ per hour for their participation.

Stimuli and Apparatus

We used simple two-dimensional geometrical objects as stimuli. Stimuli were created in the RGB color space, using MatLab (MathWorks Inc., Natick, MA). Stimuli were presented using E-Prime (Psychology Software Tools, Inc., Pittsburgh, PA) on a SVGA monitor (refresh rate: 85 Hz) at a distance of 80 cm from the participants. The stimuli measured 3 degrees of visual angle. They were either red, green, or blue; their form was a circle, a triangle or a square; and lastly they moved outward from the centre of the screen at an angle of 45 (right), 165 (down) or 285 (left) degrees. Stimuli were presented in front of an equiluminant gray background with stimulus luminance at 20 percent and stimulus saturation at 30 percent. Stimulus luminance and saturation was purposefully kept low to make color recognition more difficult. Previous experiments indicated that at full saturation, color is recognized much faster than form or direction of movement, and for the present purpose it was important to choose three attributes which are recognized within roughly the same time frame.

Target stimuli possessed one or more of the following attributes: color blue, form of a square, and moving to the right. The presence of any single one of these attributes was sufficient to define a given stimulus as a target. Non-target stimuli did not possess any of the target attributes. They were either green or red, a triangle or a circle and moving to the left or to the bottom of the screen.

Design

50% of all stimuli presented to participants were targets. To avoid contingencies between attributes on different channels which would facilitate redundant target recognition, the stimulus distribution shown in Table 1 was based on the three rules of contingency formulated by Mordkoff and Yantis (1991). However, this setup had three instead of two feature channels, thus exponentially more combinations of features needing to be balanced. We could not control perfectly the contingencies, but we did avoid positive contingencies within target attributes as well as between target and distracters attributes. This means that no facilitation of double and triple redundant targets due to existing contingencies could have occurred. However, inhibition of redundant stimulus recognition might be possible, since the non-target attributes on each channel do not occur as part of a target stimulus with the same frequency. The second non-target attribute, which will be referred to as ‘foil’, is associated with non-target stimuli more frequently than the first. This was done to avoid facilitatory contingencies. Comparing foil and non-foil target-present trials gives us the means of testing the impact of

negative contingencies (see results). Overall, none of the contingencies favored redundant target trials, reducing the chances of observing coactivation effects.

Insert Table 1 here

Procedure

The experiment consisted of 16 blocks with 60 trials per block for a total of 960 trials. Eight of the participants participated in two sessions of the same experiment, performing 32 instead of 16 blocks, with a total of 1920 trials. Stimulus distribution did not vary between blocks, but the order of trials was randomized. Participants had the possibility to take a break between blocks, and were encouraged to do so explicitly after eight blocks were completed.

The triple redundant target (target with all three target attributes present) was presented 48 times per participant. The six stimuli in double redundant conditions (any two target attributes present, plus one of two possible distracters on the third channel) were presented 48 times each. The three stimuli with only one target attribute were also presented 48 times each per participant. The difference in frequency of presentation was necessary in order to have a balanced stimulus distribution that avoided facilitatory contingencies (see above). Non-target stimuli were presented 480 times per participant

Each trial started with the presentation of a blank screen for 494 ms. The blank screen was followed by a fixation point for 694 ms. The stimulus was then presented for a total of 823 ms during 9 frames, the first positioned at the screen center for 47 ms. The subsequent 7 frames lasted 47 ms, and gradually displaced the stimulus in the required direction of movement. The last frame lasted 447 ms, showing the stimulus at its final destination 3.78 degrees of visual angle from the screen center. Finally, a feedback slide appeared for 753 ms.

For the reasons mentioned in the introduction, we used a Go-NoGo experimental paradigm. Participants were required to press the SPACE key on a keyboard as soon as they recognized a target stimulus, and discouraged from doing so if they recognized a non-target. They were encouraged to respond as fast as possible while making as few errors as possible. Responses had to happen within a time frame of 0 to 750 ms after stimulus onset.

Participants received feedback on their performance on each trial. Feedback on false responses was accompanied by a 700 Hz sound. Fast and correct performance was further encouraged by a system of points: participants were encouraged to try for the best score. Participants received 30 points for hits and correct rejections, 50 for particularly fast hits (under 300 ms), and -350 points for false alarms and misses. At the end of each block participants were given their cumulative score.

Results

Participants mastered the task very well, with an average of less than 1.4 % misses (13.03 of 960 trials per participant on average) and 2.5 % false alarms (FA; 23.78 of 960 trials per participant on average). There was no significant difference between error rates for participants which passed 16 or 32 blocks (Miss: $t(22) = -0.97$, $p = 0.34$; FA: $t(22) = -1.09$, $p = 0.29$). Neither was there a significant difference between mean response times for the two groups and seven conditions ($t(166) = 0.16$, $p = 0.87$). Therefore, we may conclude that additional training had no significant effect. All further analysis will not take the different number of blocks into account.

Trials where participants responded faster than 205 ms were excluded from analysis, as these were considered anticipatory responses (a total of four trials were eliminated). For all further analysis, only correct Go-trials will be used. While maintaining a very high performance rate, participants also responded very rapidly: valid response times to a target could be as fast as 221 ms. Participants maintained a mean response time of 398 ms across conditions, with a standard deviation (std) of 82 ms. Response times varied greatly, however, across conditions, and even more across participants. The mean response time for the triple redundant condition was 361 ms across participants, with a std of 56 ms. For double redundant conditions, the mean RT was slightly slower, at 389 ms, with a std of 74 ms. Finally, the mean response time for single-target conditions was 431 ms (std: 93 ms).

Difference in negative contingencies

It is possible that performance was influenced by negative crosstalk due to the stimulus distribution (see Table 1). To empirically test this influence on participant performance, we compared error rates and mean response times in the double redundant conditions, splitting them according to the type of non-target attribute (foil or non-foil) they possessed. These conditions contained only one non-target attribute, allowing for a direct comparison. The double redundant conditions were those where the target possessed two target attributes, either the correct color and the correct form (cf), the correct color and direction of movement (cd), or the correct form and direction of movement (fd).

Participants' error rates did not differ between types of non-target for any of the three double redundant conditions (cf: $t(46) = -0.23$, $p = 0.41$; cd: $t(46) = -0.91$, $p = 0.18$; fd: $t(46) = -0.56$, $p = 0.29$). The same holds true for mean response times: they did not differ significantly between types of non-targets for any of the double redundant conditions (cf: $t(46) = -0.65$, $p = 0.26$; cd: $t(46) = -1.66$, $p = 0.052$; fd: $t(46) = -0.26$, $p = 0.40$). We can therefore conclude that negative contingencies do not affect response times, at least in the double redundant conditions. Hereafter, no distinction will be made between different non-target types in the double redundant conditions.

Redundancy gain

Figure 1 presents the mean response times to a target in conditions where a single target attribute was presented (color only: c, form only: f, or direction of motion only: d), in conditions where two target attributes were presented (conditions cf, cd, or fd) and where all three target attributes were presented (condition cfd). As can be seen in Figure 1, there are benefits of redundant targets at the level of the mean RT, both for double redundant (2Red) over single-attribute conditions (1Red) and for triple redundant (3Red) over double redundant conditions.

Insert Figure 1 here

Insert Figure 2 here

Figure 2 shows the cumulative response time distributions of one representative participant, for all three single-target conditions, as well as for all three conditions where two target attributes were presented simultaneously. The probability of responding at time t or faster is plotted as a function of time.

Due to the large inter-participant differences in overall response time distributions, systematic changes in response times depending on condition are hard to see if analysis is done over all participants. Therefore, to test for redundancy gain, we did not test for statistically significant differences between distributions with multiple t-tests, as did Miller (1982) and Mordkoff and Yantis (1991, 1993), among others. Instead, we assessed the difference between two cumulative distributions of response times for different conditions from one participant using a two-sample Anderson-Darling test of cumulative distributions (Anderson and Darling, 1952) at a level of significance of .05. This test was chosen as it is more sensitive to differences at the extreme ends of distributions than the more well-known Kolmogorov-Smirnoff test (Engmann & Cousineau, submitted).

Response time distributions to double redundant stimuli with color and form present were significantly faster than RTs for color only for almost all participants (21 out of 24), and significantly faster than the RTs for form only (17/24 participants). The same holds for double redundant stimuli with color and direction of movement present (24/24 and 15/24 participants respectively) and for form and direction of movement (23/24 and 20/24 participants respectively). The upper part of Table 2 recapitulates the results.

Insert Table 2 here

Figure 3 shows the cumulative response time distributions for a representative participant of the three double redundant conditions as well as when all three target attributes are present. Since participants' reaction times were already fast for the double redundant conditions, an increase in RT becomes more difficult to observe for triple redundant targets. Even though a general trend was visible, and most participants responded faster for triple redundant stimuli than for any double redundant stimuli, the difference between fd and cfd conditions was significant for only 7 participants (fd: 7/24 participants). However, in the other two conditions (cd and cf), the proportion of participants that have a significantly faster distribution in the triple redundant condition is greater (cd: 18/24; cf: 24/24). Table 2, bottom part, gives the details for all the double redundant target distributions relative to the triple redundant target condition.

Insert Figure 3 here

These results show that it is indeed possible to profit from three redundant attributes for object recognition, even though the benefit of a third attribute is not always significant across participants.

Excluding Race Models

After having found that redundancy gain exists for three redundant attributes within the same modality, we now need to distinguish between the possible causes of this gain. Since facilitation by crosstalk was excluded *a priori* (see method), this leaves coactivation and statistical facilitation as possible explanations. Since statistical facilitation has an upper limit to the amount of gain it can explain (see introduction), we will use this upper limit to differentiate between these two causes of gain.

The Miller Inequality is commonly used, as it gives a definite upper limit to the performance of race models with two channels, and is therefore a very efficient criterion of exclusion for race models on any task with two redundant targets. However, one cannot estimate the degree of dependence between channels, as it is not directly observable. Hence a generalization of the Miller Inequality to three or more channels poses a problem: The degree of dependence between any two channels needs to be subtracted from the sum of the response time distributions of the separate channels, but then the degree of dependence of the three channels would have to be added again. Both values are positive but of unknown size, making it impossible to determine in which direction the extension of the Miller Inequality would tend.

An alternative to the Miller Inequality was proposed by Townsend and Nozawa (1995; a similar bound was proposed by Mordkoff and Yantis, 1991, p. 535). It is based upon survivor functions (one minus the cumulative distribution)

of response times instead of cumulative distribution functions. The upper limit to race model performance with more than one channel is given by the survivor function of the product of the survivor functions of each channel:

$$Pr(\mathbf{RT}_{cdf} < t) = 1 - \prod_{i \in \{c,d,f\}} 1 - Pr(\mathbf{RT}_i < t) \quad (1)$$

where \mathbf{RT}_{cdf} is a response time in the triple redundant condition and i indexes the three conditions with only one target attribute. If the observed response time distribution in a redundant target task is significantly faster than predicted by this boundary, race models as the sole explanation of redundancy gain are rejected. The Townsend Bound can be calculated for any number of channels. We again used the two-sample AD test to test for significant differences between the redundant target distribution and the Townsend Bound calculated from the corresponding single-target RT distributions.

Figure 4 shows the single-target distributions of a representative participant, as well as the Townsend Bounds and the actual response time distributions for each of the three double redundant conditions. The AD test compares the dotted line to the full line in this figure.

Insert Figure 4 here

For double redundant targets with color and form present, the RT distributions were significantly faster than the Townsend Bound calculated from the RT distributions of single-target color-only and form-only targets for two participants only ($AD = 2.87; 4.23$; critical value $AD = 2.49$ for $\alpha = .05$). For color and direction of movement, four participants' RT distributions were significantly faster than the corresponding Townsend Bound ($AD = 3.09; 3.43; 3.68; 5.09$); and for form and direction of movement twelve participants' RT distributions were significantly faster than the corresponding Townsend Bound ($AD = 7.18; 4.92; 4.39; 5.52; 7.07; 10.03; 2.99; 12.43; 5.16; 37.25; 2.85; 8.44$).

To test for coactivation in the triple redundant target condition, the Townsend Bound was calculated from all three single-target RT distributions. Figure 5 shows the three single-target distributions of a representative participant, the Townsend Bound and the actual distribution for the triple redundant condition. Eleven out of 24 participants performed significantly faster than predicted by the Townsend Bound on triple redundant target trials ($AD = 5.42; 4.64; 6.43; 4.23; 3.81; 3.34; 8.56; 13.25; 2.50; 5.13; 6.89$).

Insert Figure 5 here

Excluding race models for gain from double to triple redundancy

The Townsend Bound for triple redundant targets, as formulated above, is based on the RT distributions of single-target trials. It is able to assert that the interaction between three single target attributes is more than statistical facilitation. This opens the possibility that the third attribute contributes only within the range of statistical facilitation. In order to show that statistical facilitation cannot be responsible for specifically the amount of gain obtained from an additional third attribute, we need to adapt the Townsend Bound. We can calculate the Townsend Bound for triple redundancy for the RT distribution of any of the three double redundant conditions plus the RT distribution of the respectively missing single-target condition. We then obtain the final Townsend Bound by taking the maximum of the three above-mentioned values at each time point.

$$Pr(RT_{cdf} < t) = \max[1 - (1 - Pr(RT_{cf} < t)) \times (1 - (Pr(RT_d < t))), \\ 1 - (1 - Pr(RT_{cd} < t)) \times (1 - (Pr(RT_f < t))), \\ 1 - (1 - Pr(RT_{fd} < t)) \times (1 - (Pr(RT_c < t)))] \quad (2)$$

This yields the most liberal evaluation of performance if any of the three target attributes contributes only by statistical facilitation as a third redundant attribute. Hence, exceeding this limit can only be achieved if all three target attributes contribute significantly to the amount of redundancy gain.

The Townsend Bound for triple over double redundancy could have been calculated from the product of the survivor functions of the three double redundant

conditions, analogous to equation (1). However, in this case, the gain contributed by each attribute would be included twice (once in each of the two double redundant conditions it is part of), thereby obtaining an upper limit which would definitely exceed performance of a combination of coactivation for two and statistical facilitation for the third attribute. Equation (2) ensures that each attribute contributes only once, while still ensuring the best possible performance under the assumption that statistical facilitation is responsible for the gain attributed to the third target attribute.

Two participants performed significantly faster on triple redundant trials than predicted by the triple over double redundancy Townsend Bound ($AD = 2.9; 2.5$). Figure 6 shows the cumulative RT distributions for double redundant and triple redundant conditions for each of these participants respectively, as well as the Townsend Bound obtained from the non-redundant conditions and the Townsend Bound obtained from the double redundant conditions. This shows that gain solely from a third redundant attribute cannot be accounted for by statistical facilitation. We can therefore conclude that all three target attributes must interact in some other way than statistical facilitation to contribute to a triple redundancy gain.

Insert Figure 6 here

In conclusion, we show that redundancy gain from a third attribute is indeed possible, even though all attributes come from the same modality and despite the fact that performance for double redundant targets is already very fast.

The Townsend Bound is not violated for every participant and every condition. This can have several possible causes. First, since participants responded very rapidly, even to single-target conditions, room for improvement under redundant conditions is limited. This is supported by the fact that 18 of 24 participants' triple redundant distributions tended to be above the Townsend Bound, even if the AD test did not prove this trend to be significant in 13 of 24 participants. Second, the potential for improvement under redundant conditions is maximal if single attributes are processed at approximately the same rate. If one attribute is visibly slower than another, the advantage of adding this target attribute is smaller than if attributes are processed at the same speed. As can be seen in Figure 1, the attribute color may be slightly slower than the attribute direction of movement. Lastly, gain from redundant target conditions can be outweighed by sources of noise or inhibition. Although there are several possible internal sources of noise, one should be mentioned specifically. We excluded crosstalk as an explanation of redundancy gain by avoiding contingencies that *facilitate* redundant object recognition. We could not however avoid *inhibitory* contingencies. If we assume that crosstalk between channels exists, it would work against a gain from redundancy under these experimental conditions. However, our analysis of the impact of foils versus non-targets (see above) seems to indicate that inhibitory contingencies had no or only a negligible impact on the data.

Nevertheless, the violations we did obtain are enough to refute statistical facilitation as an incomplete explanation of redundancy gain. Since we excluded crosstalk as a possible explanation by the specific experimental design we used, this leaves us with only coactivation as an option for explaining the size of the redundancy gain observed under these conditions. However, before accepting coactivation as a default explanation for redundancy gain, its limits and predictions need to be tested more vigorously.

Characteristics of the RT distributions

Before we test the coactivation model more rigorously in the next section, we examined the distributions of RTs more thoroughly in order to get estimates of the minimal response times and skew. To do so, we chose to model RT distributions using a Weibull distribution for several reasons. First, there is some evidence that it captures accurately the characteristics of an RT distribution (Logan, 1992; Rouder, Lu, Speckman, Sun & Jiang, 2005). Second, the Weibull distribution is the predicted distribution for many sampling models such as the race model (Cousineau, Goodman, & Shiffrin, 2002; Galambos, 1972). Lastly, the Weibull distribution is very flexible, as it can resemble an exponential or a normal distribution depending on its shape parameter (Heathcote, Brown and Cousineau, 2004). Due to this flexibility, the Weibull distribution is often used to model response time distributions without theoretical commitment.

Estimating the minimal response times by fitting a Weibull to the raw data instead of taking the fastest response times per participant and condition increases

the reliability of the estimation, as the estimated Weibull minimum is based upon a whole set of data points instead of just a single one. A single data point is much more subject to noise, such as anticipatory responses or moments of inattention (Hirose & Lai, 1997).

The Weibull cumulative distribution function is given by:

$$F(t) = 1 - e^{-\left(\frac{t-\alpha}{\beta}\right)^\gamma} \quad (3)$$

where α gives the starting point of the distribution, i.e. the minimal response time; β is the scale parameter, which indicates the range, or spread of the data (comparable to the σ parameter of normal distributions); and γ represents the degree of symmetry, i.e. the length of the tail and the degree to which the data lean to one side or the other. For $\gamma = 1$, the distribution is exponential; for $\gamma = 3.6$, the distribution is close to symmetrical.

insert Figure 7 here

Figure 7 presents the mean estimated α as well as the means for β and γ (error bars are computed after removing between-participant differences; Loftus & Masson, 1994; Cousineau, 2005, Morey, 2007). We have partitioned the participants to separate those that exceeded the Townsend Bound (left bars for

redundant conditions) from those that did not (right bars) to check if there is a qualitatively different pattern of results between them.

As can be seen, the same trends are visible for the two partitions: decreasing α and β with increasing redundancy and increasing γ with increasing redundancy. The two partitions seem to differ only with respect to the parameter β in two of the double redundant conditions: participants that did not exceed the Townsend Bound apparently have more variable RTs (β is the spread parameter) which may explain at least in part why the Anderson-Darling test was not significant. In the following (and top of Figure 8), we no longer partition the participants.

Regarding the minimal response times, all participants had fairly large α s in the non-redundant conditions (mean: 311 ms, std: 3 ms; after removing between-participant differences; Loftus & Masson, 1994; Cousineau, 2005), smaller in the double redundant conditions (mean: 283 ms, std: 3 ms), and even smaller in the triple redundant condition (mean: 274 ms, std: 4 ms). These differences were significant (One-way ANOVA: $F(2,46) = 12.35$, $p < 0.001$). This result alone gives us an additional argument for refuting statistical facilitation as a cause for redundancy gain, since minimal response times for redundant conditions would necessarily have to be drawn from the RT distributions of non-redundant conditions in the case of a race model, which is clearly not what is happening here. Regarding the parameter β , the decreasing trend with increasing redundancy is also significant ($F(2,46) = 8.14$, $p < 0.001$). Finally, the opposite trend is significant for

the parameter γ ($F(2,46) = 28.82$, $p < 0.001$). These last two trends seem to be remarkably linear.

Predictions of the coactivation model

Now that we have a better characterization of the results, we can examine the coactivation model more attentively. We have rejected the race models for two reasons, the RT distributions exceeded the Townsend Bound and the minimum response times were not constant. The argument often found is that if the RTE is not caused by a race model and cannot be due to crosstalk (owing to the Mordkoff and Yantis controls), then it ought to be a proof of coactivation. However, before we accept this as an argument, we need positive evidence favoring a coactivation model. In the following, we will show that such evidence is not present in the data.

Coactivation models account nicely for the reduction in minimum response time with redundancy. For instance, suppose that the threshold size to trigger a response is k . In a double redundant condition, half of these might come from one channel, the other half from the other channel. Having to sample less evidence per channel, the minimum response time would be lowered. As a rough approximation, the reduction in α should mimic the reduction in the number of evidences per channel, that is, be inversely proportional to redundancy. This prediction is not incompatible with the results seen in Figure 7(a).

However, coactivation model makes one critical prediction. The symmetry (measured using the Fisher skew or with the parameter γ) in the RT distribution

should remain stable between the triple redundant and non-redundant conditions, and decrease significantly for the double redundant condition: The model predicts almost symmetrical RT distributions, for all redundancy conditions except double redundancy and turns to perfect symmetry as redundancy increases beyond three. Expressed in terms of the parameter γ , after an initial decrease, γ rapidly tends towards 3.6. This argument was first established using Monte Carlo simulations (see below and appendix) and confirmed using numerical integration techniques (next section).

The Monte Carlo simulations also confirmed that under conditions of coactivation, α , the parameter giving the starting point of the RTs was repeatedly and reliably lower for the triple redundant condition than for the non-redundant one. The scale-parameter β also decreased for triple redundant conditions, meaning that the RT distribution became narrower. Whereas the empirical α and β behave as predicted by a coactivation model, the symmetry parameter did not: first, γ was smaller in the triple redundant condition than in the non-redundant condition. Second, the distributions were far from symmetrical, although they tended towards symmetry.

insert Figure 8 here

Figure 8 directly compares the best-fitting parameters for all participants (left side) and the parameter predictions of the model (right side). As can be seen, reality and prediction concord for parameters α and β , not so however for γ .

As seen in Figure 7(c), all participants, whether they exceeded the Townsend Bound or not (using the AD test), showed the same pattern of change in γ . However, this pattern is not compatible with what the coactivation model predicts. Instead of decreasing and then increasing for redundant conditions, γ was larger for double redundant conditions than for non-redundant conditions, and even larger for triple redundant conditions. The response time distributions of participants become more and more symmetrical as the number of redundant target attributes increases. This is neither what the coactivation model predicts, nor what one would intuitively expect for a speeded response task.

The cascade race model

In the following, it will be useful to define the operator $\&_{k:R}(F)$ which returns the cumulative distribution function of the k^{th} fastest signal out of R given that each signal has a distribution function F . If the decision threshold $k = 1$, the fastest of R channels is given by:

$$\&_{1:R}(F)(t) = 1 - (1 - F(t))^R \quad (4)$$

because $Pr(T_{(1)} < t) = 1 - Pr(T_{(1)} > t \& T_{(2)} > t \& \dots \& T_{(R)} > t)$ where $T_{(i)}$ is the i^{th} fastest time (for $i = 1 \dots R$). This is precisely the Townsend Bound and therefore no race model with a threshold of $k = 1$ can predict performance above it.

In addition, if the distribution of a single channel F is member of the class of power distributions (PD; Gumbel, 1958), that is, if its left tail follows a power curve such that

$$\lim_{t \downarrow \alpha} \frac{F(t + \alpha)}{t^\gamma} \quad (5)$$

is a constant, in which α is the lower bound of the distribution and γ is any exponent (Gnedenko, 1943; Galambos, 1978), then for large R , the distribution tends to a Weibull distribution

$$\&_{1:R}(F) \rightarrow Weibull(\alpha, \beta, \gamma) \quad (6)$$

where the shape parameter of the attractor distribution γ is the shape and the lower bound α is the lower bound of the individual channels (Cousineau, Goodman, & Shiffrin, 2002). The convergence to the attractor distribution is usually fast so that

even if there are only 2 or 3 channels, the resulting distribution should be very similar to a Weibull distribution. Hence, the race model with $F \in \text{PD}$ predicts that shape is constant irrespective of redundancy.

If the distribution of the individual evidence is not a member of the PD class, the attractor distribution is not the Weibull distribution (it can be a double-exponential distribution, a Gumbel distribution or it may not have an attractor distribution).

In the following, we examine the scenario where the channels send activation following a cascade model (McClelland, 1979, Ashby, 1982). In this model, stronger signals (and presumably faster) will be reinforced more as they travel through a few layers so that initially strong signals will get stronger. This cascade model can be implemented in many different ways (e.g. as a succession of Kohonen self-organizing maps; Kohonen, 1984). Ulrich and Miller (1993) showed that the attractor distribution (under some general assumptions) is the lognormal distribution whose cumulative distribution function is given by

$$F(t) = \Phi \left(\frac{\log(t) - \mu}{\sigma} \right) \quad (7)$$

in which Φ is the cumulative distribution function of the normal distribution and μ and σ are shift and spread parameters.

Because the lognormal distribution has a long left tail, the shape of the distribution of the fastest among R channels will change: As R increases, the

probability of picking fast evidences from within the left tail increases and this increase changes the shape of the distribution, making it more symmetrical. Assuming we fit a Weibull distribution to this distribution, we could derive the estimated γ of this Weibull distribution by applying equation 4 and then finding the expected Fisher skew. For example, if $\mu = 2.5$ and $\sigma = 0.35$, the estimated γ would be 1.5, 1.9 and 2.1 for single, double and triple redundancy conditions respectively. Likewise, the estimated α would be 6.3, 5.1 and 4.6 in the same conditions. These results fit nicely the mean estimated parameters from the experiment. However, as said earlier, this model does not predict coactivation.

The basic version of the race model (illustrated in Figure 9, top) is extended by assuming that the decision threshold is a free parameter greater than one. In this case, it might be necessary to receive multiple evidences from each channel (e.g. when $k > R$). This can be done in the two possible ways to be reviewed next. Because evidences can fill the decision accumulator from any channel, this results in a coactivation model.

The serial coactivation model

This is the model generally known by the name of “coactivation model”. In this model, a channel can send multiple evidences separated in time by a delay t whose distribution function is labeled F in the following. This model is illustrated in Figure 9, bottom left.

Insert Figure 9 about here

The time to sample k evidences, $\mathcal{E}_{k:R}(F)$, is a mixture of all the ways k evidences can be sampled. For example, if $k = 2$, two evidences could be received from the first channel and none from all other channels, or 1 from one channel, 1 from a second channel, and none from the others. In conditions of triple redundancy, the first scenario could occur in three different ways and the second, in six different ways. In the first scenario, the time to sample two evidence is given by a second-order convolution ($F * F$; Luce, 1986, Cramér, 1946).

As soon as $k > 1$, it is easy to see why the coactivation model predicts a massive violation of the Townsend Bound. With one cue, the decision time requires k evidences, which are received serially. Mean response time will be k times the mean response time of a single evidence; with two cues, each channel only needs to send half as many evidences to reach the decision threshold, so that the mean response time is half the above (this is not the only way that the response threshold can be reached, but it is certainly the most frequent). The violation of the Townsend Bound predicted by this model is so large that the predicted distribution for the double-redundant conditions exceeds the triple-redundant Townsend Bound.

We also examined the changes in symmetry of the predicted response time distributions as a function of redundancy. The changes were very small: If the

evidence distribution F is exponential (large positive symmetry), the degree of symmetry is a constant whereas if F is symmetrical, the degree of symmetry diminishes a little for the double redundant condition and increases afterwards, conforming the Monte Carlo of the previous section.

For these two reasons (too large of a coactivation effect and nearly no changes in symmetry with increasing redundancy), the serial coactivation model is rejected.

The parallel coactivation model

This version of the model assumes that the internal channels are redundant. Each input can convey evidence through ρ redundant channels working in parallel. In this situation, the response time distribution is the k^{th} fastest evidence from a pool of $\rho \times R$ channels. It is given by

$$\&_{k:R}(F)(t) = \int_0^t \frac{(\rho R)!}{(\rho R - k)!(k-1)!} F^{k-1}(t) \times f(t) \times (1 - F(t))^{\rho R - k} dt \quad (8)$$

(David, 1970). In the case where $k = 1$, it is easy to show that

$$\&_{k:\rho R}(F) = 1 - (1 - \&_{k:\rho}(F))^R \quad (9)$$

so that this version of the model simplifies to the basic race model seen above.

Further, either $F \in \text{PD}$ in which case $\&_{1:\rho}(F)$ is also $\in \text{PD}$ or $F \notin \text{PD}$ in which case

$\&_{1:\rho}(F)$ is not as well.

For k larger than one, the same dichotomy is valid: either $F \in \text{PD}$ so that $\&_{k:\rho R}(F) \rightarrow \text{Weibull}(\alpha, \beta, \gamma)$ (but convergence is slower) or F is not a member of the PD class, in which case $\&_{k:\rho R}(F)$ is also $\notin \text{PD}$.

It is possible to show that

$$\&_{k:\rho R}(F)(t) > \left(1 - (1 - \&_{k:\rho}(F))^R\right)(t) \quad (10)$$

that is, that pooling the activations of $\rho \times R$ channels is better than examining separately the k^{th} fastest of the attributes for each channel independently, and then taking the fastest of the R winners. Hence, the parallel model predicts a violation of the Townsend Bound. This violation is much smaller though, with the predicted response time distribution for two redundant cues at about the level of the triple-cue Townsend Bound.

When the distribution F is a member of the PD class, the symmetry is very close to constant across the redundancy conditions, the small deviations (at third decimal place for $\rho = 5$ and $k = 2$) reflecting the fact that convergence to the asymptotic shape is slower when $k > 1$.

Among the possible distributions which are not members of the PD class, we chose the lognormal distribution for the same reason as above: a channel is assumed to be a cascade of activations reaching the decision node. This model predicts that the symmetry increases with increasing number of attributes on the target, as observed in the experiment. In this model, the parameter μ is a scaling

parameter and does not change the symmetry as internal (ρ) or external (R) redundancy is increased.

Table 3 summarizes the findings for the three models explored, according to whether F is member of the PD class or not.

Insert Table 3 here

Insert Figure 10 here

Figure 10 shows the estimated symmetry as R, ρ and k are increased. There is a trade-off between k and ρ so that if both are increased, the observed symmetry stays roughly constant (middle panel).

To assess the capability of the model to capture the results, we found the best-fitting parameters by maximizing the log likelihood. For simplicity's sake, the cues are assumed to elicit the same response time distributions. We added a shift parameter to the model, T_0 . In all, it has five free parameters: μ and σ , the characteristics of one channel, k and ρ , the decision threshold and the internal

redundancy and finally, T_0 . The index of fit is shown for each subject individually in Table 4.

Insert Table 4 here

To validate the quality of the fits, we compared them to two descriptive models: one fitting three Weibull distributions, one per redundancy condition to extract the parameters α , β , and γ , for a total of nine parameters. The other descriptive model fitted three lognormal distributions with parameters μ , σ , and a shift parameter ξ in each of the three redundancy conditions, for a total of nine parameters as well.

Insert Table 5 here

The mean best-fitting parameters across participants are given in Table 5. As seen, the 5-parameter cascade race model fits the data less well than the 9-parameter 3lognormal model but 21 times out of 24 better than the 3Weibull model. Using the Bayesian Information Criterion penalty term $\log(n) \times \Delta p / 2$ where n is the number of data, p is the number of parameters, and Δp the

difference between the number of parameters in the two models, the penalty term is equal to 12.3 (Hélie, 2006), we find that the 11 out of 24 participants have a fit significantly worse than with the 3lognormal model. However, this penalty term is approximate. In comparison, the 3Weibull model is significantly worse than the 3logNormal model for 19 of the 24 data sets. When the average fit is examined, the 3Weibull model is rejected but not the cascade race model. Hence, we cannot reject the cascade race model based on best-fitting results.

In addition, the estimated shift parameter is more plausible for the cascade race model than it is for the 3lognormal model (an estimate of 249.6 ms, nearly comparable to the estimated shift of the 3Weibull model, 278.8 ms, much different from the 3lognormal model estimate of 185.2 ms).

Some conclusions related to model fitting

The fact that the 3Weibull model has reasonable best-fitting parameters but nevertheless offers the poorest fit of the data suggests that there is something inherently wrong with this model. The major characteristic that differentiates a Weibull distribution from a lognormal distribution when the data are positively skewed is the left tail: the Weibull distribution has a very short left tail. Hence, even though the Weibull model is convenient (Rouder et al, 2005, Cousineau and Shiffrin, 2004), has nice asymptotic properties in connection with the PD class (Cousineau, Goodman and Shiffrin, 2002, Galambos, 1978) and can be fitted efficiently and without bias (Cousineau, 2009 (a); Cousineau, 2009 (b)), its

relevancy to response time analyses, at least for the redundancy paradigm, must be questioned seriously.

If we accept the assumption that redundant cues are used simultaneously to make a decision about the presence of a target, then we have shown that the decision threshold has to be larger than one (violation of the Townsend Bound) and that multiple cues must be sent in parallel (the serial race model predicts too important a violation of the Townsend Bound). Finally, we have shown that the parallel model cannot be fed by inputs whose distributions are in the PD class (or else the symmetry is a constant). The failure of the serial coactivation model is as informative as the success of the parallel version of the model as it constrains considerably the development of alternative models.

We have suggested that a channel could be caused by a cascade of activation in which case the lognormal occurs as the asymptotic model (the attractor distribution). However, we have not tested this suggestion with other distributions that are not members of the PD class.

The joint findings of violation of the Townsend Bound and of a continuous increase in symmetry turn out to compose a major challenge as very few models are able to produce these results simultaneously.

Discussion

This study is the first to find a triple redundancy gain from attributes within a single modality. Not all participants showed a triple redundancy gain, and for those that did, not all overcame the threshold given by the Townsend Bound. This, as well as the lack of other studies with similar findings, is due to the difficulty of creating the right conditions under which such a gain might be observed. There are several factors to be considered.

First, the choice of attributes is important. Attributes need to be processed at approximately the same speed, since an attribute that is processed much slower than another will not be able to contribute substantially to an RTE. It is not possible to know a priori which attributes are processed at the same speed, and also, different attributes are processed at different speeds at different stages of the processing pathway. For example, detection of movement is very fast, whereas recognition of color is faster than recognition of form (Kandel et al., 2000). Due to the difficulty of measuring, or of trying to equalize processing speed at different levels of the processing pathway, the choice of attributes for this study was made based on pilot experiments testing the overall reaction times to a series of attributes. From among orientation, spatial frequency, size, letters, direction of movement, color, and form - with varying degrees of saturation and luminosity to modify processing speed of certain attributes – the latter three proved to elicit approximately the same reaction times for most participants, under conditions of low luminosity and saturation. The argument that attributes need to be processed at

approximately the same speed to be able to observe maximal triple redundancy gain is supported by the results of the present study: Redundancy gain was found for those participants that did indeed process all three single attributes at the same speed, and more specifically also processed double redundant targets at roughly the same speed.

Second, participants had different patterns of responses. Some reacted slower to certain attributes than to others, thus making a triple redundancy gain more difficult to observe. Others had very narrow RT distributions, i.e. very steep cumulative RT distributions, thus making the distance between two distributions very hard to detect. Others again were simply extremely fast, even in the non-redundant condition, thus leaving very little room for an improvement in performance under redundant conditions.

Third, noise from external as well as internal factors can mask any redundancy gain. The participant might be concentrating mainly on one attribute, or show slight hesitation in their motor response, or there might be some noise added to the transmission of the signal at any stage of the processing pathway. Since there is already very little room for improvement from double to triple redundancy, it would not take much to mask a triple RTE.

Finally, when excluding crosstalk as a facilitating influence on the RTE, contingencies between attributes that would inhibit an RTE were permitted. Since the possibility that crosstalk exists cannot be excluded, this inhibition could well

have masked an RTE, even though we checked that this effect is not important (comparing non-target and foil stimuli).

Despite these difficulties, the current results allow an important conclusion. Finding a triple redundancy gain that violates the Townsend Bound is possible, and it cannot be explained by one-level race models.

This study proposes the cascade race model as a novel explanation of the RTE. The cascade race model is basically an extension of the already existing race model. It shows that coactivation can explain the present results if and only if there is internal redundancy (channels sending information multiple times in parallel). In addition, the internal channels must have a distribution which is not part of the Power distribution class. Those negative findings constrain possible future modeling efforts of the RTE.

Miller (1982) found that when a single-letter condition (no distracter on the second channel) was added to the letter search task, the Miller Inequality is not violated. Grice and colleagues (1984) and Van der Heijden and colleagues (1984) both confirmed the importance of distracter presence for the violation of the Inequality. One possible explanation is that due to resource competition, single-target trials with distracter presence are much slower than single-target trials without any distracters. This gain from attentional focus might be sufficient to balance the advantage on redundant signal trials. The cascade race model simulates this effect of attentional focus by increasing the rate of information transmission in the internal channels. If the attributes are processed by Kohonen

Self-Organizing Maps (Kohonen, 1984), fewer information leads to faster stabilization of the Map.

The cascade race model is not the first model that reconciles race models with a violation of the Miller or Townsend Bound. Earlier attempts to reconcile race models with violations of the Miller Inequality are the Parallel Grains Model (PGM; Miller & Ulrich, 2003) and the Parallel Race Model (PRM; Cousineau, 2004). They are separate activation models representing stimuli by a certain number of activation grains (redundant evidence), which travel towards a decision center with varying speed. Once a sufficient number of grains have arrived, the decision threshold is reached and a response can be initiated. This enables them to modelize time, which other race models cannot. The PGM can be extended to belong to the class of coactivation models by assuming a separate pool of grains for each input channel, which then feed into a unique decision center. It is able to modelize response time violation of the Miller Inequality as well as the temporal dependency of the RTE on stimulus onset asynchrony and its dependency on signal intensity. However, RTE dependence on spatial factors such as distance between two visual target and effects of grouping was not tested, and one would not expect the PGM to explain these easily. The cascade race model is at an advantage over the PGM and PRM due to its greater simplicity: It is able to explain the results on redundancy gain found in literature employing less parameters.

There has been an extensive discussion concerning the stage along the processing pathway which is responsible for violating the Miller Inequality. Since this violation was mainly taken as evidence for coactivation, various types of coactivation models have been proposed, with coactivation happening at different stages of the processing pathway. Fidell (1970) and Miller (1982) both assume coactivation takes place at a very early stage of processing, the detection level (activation from both channels is pooled in order to detect a signal in either modality). However, other variants of coactivation models have been proposed. The Logogen model (Morton, 1969) predicts performance on bimodal word recognition tasks, and postulates coactivation during the recognition process (activation from all channels which have signaled the detection of a signal is pooled to enable identification). Keele (1973) extended the Logogen model to include coactivation at the decision level (signals on different channels are recognized separately but feed into a common pool to decide the appropriate response). A model proposed by Logan (1980) predicts performance during response competition (e.g., the Stroop task), also with coactivation at the decision level. A model of visual search with coactivation at the decision level has also been proposed (Eriksen & Schulz, 1979). But even if all processing up to the decision level is based upon separate activation, response times can still profit from coactivation at the response initiation level or even at the motor stage. Input even from unmonitored modalities increases the general state of arousal, thereby facilitating response initiation (Nickerson, 1973). Differences in reaction times between responses with the right or left hand (only right hand responses showed

enough redundancy gain for coactivation) indicate that at least part of the facilitation for redundant targets arises at a motor level (Diederich & Colonius, 1987).

The cascade race model may suggest a way to explain these results. First, the channels in the cascade race model are not simply conducting activation. Indeed, normal spiking is believed to follow a Poisson distribution, which is a member of the Power distribution class (Tsodyks & Markram, 1997). If channels were only transmitting activations, they would also be a member of that class, as would the decision from a parallel race model (Cousineau, Goodman & Shiffrin, 2002). Hence, the channels must be transducers that actively manipulate the information. Second, the response times produced by the cascade race model resemble a lognormal distribution (and the 3LogNormal descriptive model fitted the data as well as the cascade race model). Hence, one possibility is that the channel times are not lognormal (as used in the previous section) but rather themselves the output of other coactivation stages occurring at earlier stages. Although it is not possible with fitting techniques to identify them, it would nicely account for the above-cited literature.

In conclusion, the results of this study show that redundancy gain from a third attribute within one modality is possible, that this redundancy gain is not compatible with simple race models, with crosstalk models, or with coactivation models. All the results of this study can however be explained by a cascade race model. The cascade race model is a novel and unifying approach to previous

research on redundancy gain and its predictions at the level of the RT distributions were confirmed.

Acknowledgements

The authors would like to thank Dominic Charbonneau and Laurence Richard for their help and Sophie Callies, Étienne Dumesnil and Laurence Morissette for comments on a previous version of the manuscript. This research was supported by the *Conseil pour la recherche en sciences naturelles et en génie du Canada*, the *German Academic Exchange Service (DAAD)* and the *Friedrich-Ebert-Stiftung*.

References

- Anderson, T. W., & Darling, D. A. (1952). Asymptotic theory of certain goodness-of-fit criteria based on stochastic processes. *Annals of Mathematical Statistics*, 23, 193-212.
- Ashby, F. G. (1982). Testing the assumptions of exponential, additive reaction time models. *Memory & Cognition*, 10, 125-134.
- Cramer, H. (1946). *Mathematical Methods of Statistics*. Princeton, NJ: Princeton University Press.
- Colonus, H., & Diederich, A. (2004). Multisensory Interaction in Saccadic Reaction Time: A Time-Window-of-Integration Model. *Journal of Cognitive Neuroscience*, 16(6), 1000-1009.
- Cousineau, D. (2004). Merging race models and adaptive networks: A parallel race network. *Psychonomic Bulletin & Review*, 11, 807-825.
- Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, 1, 75-78.
- Cousineau, D. (2009a). Nearly unbiased estimates of the three-parameter Weibull distribution with greater efficiency than the iterative likelihood method.

British Journal of Mathematical and Statistical Psychology. 62(1), 167-191.

Cousineau, D. (2009b). Fitting the three-parameter Weibull distribution: Review and evaluation of existing and new methods. *IEEE Transactions on Dielectrics and Electrical Insulation*. 16(1), 281-288.

Cousineau, D., Goodman, V. W., & Shiffrin, R. M. (2002). Extending Statistics of Extremes to Distributions Varying in Position and Scale and the Implications for Race Models. *Journal of Mathematical Psychology*, 46, 431-454.

Cousineau, D., & Shiffrin, R. M. (2004). Termination of a visual search with large display size effects. *Spatial Vision*, 17, 327-352.

David, H. A. (1970). *Order Statistics*. New York: John Wiley and Sons.

Diederich, A. (1995). Intersensory facilitation of reaction time: Evaluation of counter and diffusion coactivation models. *Journal of Mathematical Psychology* 39, 197-215.

Diederich, A., & Colonius, H. (1987). Intersensory facilitation in the motor component? A reaction time analysis. *Psychological Research*, 49, 23-29.

- Diederich, A., & Colonius, H. (2004). Bimodal and trimodal multisensory enhancement of reaction time: Effects of stimulus onset and intensity. *Perception and Psychophysics*, 66, 1388-1404.
- Engmann, S., & Cousineau, D. (submitted). Comparing distributions: the two-sample Anderson-Darling test as an alternative to the Kolmogorov-Smirnoff test.
- Eriksen, C., & Schulz, D. (1979). Information processing in visual search: A continuous flow conception and experimental results: *Perception and Psychophysics*, 25, 249-263.
- Feintuch, U., & Cohen, A. (2002). Visual attention and co-activation of response decisions for features from different dimensions. *Psychological Science*, 13, 361-369.
- Fidell, S. (1970). Sensory Function in Multimodal Signal Detection. *Journal of the Acoustical Society of America*, 47(4), 1009-1015.
- Galambos, J. (1978). *The Asymptotic Theory of Extreme Order Statistics*. New York: John Wiley and Sons.
- Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Annals of Mathematics*, 44, 423-453.

- Grice, G. R., Canham, L., & Gwynne, J. W. (1984). Absence of redundant-signals effect in a reaction time task with divided attention. *Perception and Psychophysics*, 36, 565-570.
- Grice, G. R., & Reed, J. M. (1992). What makes targets redundant? *Perception and Psychophysics*, 51(5), 437-442.
- Gumbel, E. J. (1958). *The Statistics of Extremes*. New York: Columbia University Press.
- Heathcote, A., Brown, S., & Cousineau, D. (2004). QMPE: Estimating Lognormal, Wald and Weibull RT distributions with a parameter dependent lower bound. *Behavior Research Methods, Instruments, & Computers*, 36, 277-290.
- Hirose, H., & Lai, T. L. (1997) Inference from grouped data in three-parameter Weibull models with applications to breakdown-voltage experiments, *Technometrics*, 39, 199, 210.
- Hélie, S. (2006). An introduction to model selections: Tools and algorithms. *Tutorials in Quantitative Methods for Psychology*, 2, 1-10.
- Huber, D. E., & Cousineau, D. (2004). A race model of perceptual forced choice reaction time. *Proceedings of the 25th Annual Meeting of the Cognitive Science Society, Boston*.

- Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (2000). *Principles of Neural Science*. New York: McGraw-Hill.
- Keele, S. (1973). *Attention and human performance*. Pacific Palisades, CA: Goodyear.
- Kinchla, R. A. (1974). Detecting target elements in multielement arrays: A confusability model. *Perception and Psychophysics*, 15, 149-158.
- Kinchla, R. A., & Collyer, C. E. (1974). Detecting a target letter in briefly presented arrays: A confidence rating analysis in terms of a weighted additive effects model. *Perception and Psychophysics*, 16(1), 117-122.
- Kohonen, T. (1984). *Self-organization and associative memory*. Berlin: Springer Verlag.
- Krummenacher, J., Müller, H. J., & Heller, D. (2001). Visual search for dimensionally redundant pop-out targets: Evidence for parallel-coactive processing of dimensions. *Perception and Psychophysics*, 63, 901-917.
- Krummenacher, J., Müller, H. J., & Heller, D. (2002). Visual Search for Dimensionally Redundant Pop-Out Targets: Parallel–Coactive Processing of Dimensions Is Location Specific. *Journal of Experimental Psychology: Human Perception and Performance*, 28(6), 1303–1322.

- Logan, G. D. (1980). Attention and automaticity in Stroop and priming tasks: Theory and data. *Cognitive Psychology*, 12, 523-553.
- Logan, G. D. (1992). Shapes of reaction-time distributions and shapes of learning curves: a test of the instance theory of automaticity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, 883-914.
- Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, 1, 476-490.
- Luce, R. D. (1986). Uniqueness and Homogeneity of Ordered Relational Structures. *Journal of Mathematical Psychology*, 30, 391-415.
- Marzi, C. A., Smania, N., Martini, M. C., Gambina, G., Tomelleri, G., Palamara, A., Alessandrini, F., & Prior, M. (1996). Implicit redundant-target effects in visual extinction. *Neuropsychologia*, 34(1), 9-22.
- McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, 86, 287-330.
- Miller, J. (1978). Multidimensional same-different judgments: Evidence against independent comparison of dimensions. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 411-422.

- Miller, J. (1981). Global precedence in attention and decision. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 1161-1174.
- Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, 14, 247-279.
- Miller, J., & Ulrich, R. (2003). Simple reaction time and statistical facilitation: A parallel grains model. *Cognitive Psychology*, 46, 101-151.
- Milner, A. D., & Goodale, M. A., (1993). Visual pathways to perception and action. *Progress in Brain Research*, 95:317-337.
- Morey, R. D. (2007). Confidence Intervals from Normalized Data: A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology*, 4(2), 61-64.
- Mordkoff, J. T., & Yantis, S. (1991). An interactive race model of divided attention. *Journal of Experimental Psychology: Human Perception and Performance*, 17, 520-538.
- Mordkoff, J. T., & Yantis, S. (1993). Dividing attention between color and shape: Evidence of co-activation. *Perception and Psychophysics*, 53, 357-366.
- Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, 76, 165-178.

- Mulligan, R. M., & Shaw, M. L. (1980). Multimodal signal detection: Independent decision vs. integration. *Perception and Psychophysics*, 28, 471-478.
- Nickerson, R. S. (1973). Intersensory facilitation of reaction time: Energy summation or preparation enhancement? *Psychological Review*, 80(6), 489-509.
- Raab, D. H. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*, 24(5), 574-590.
- Rouder, J. N., Lu, J., Speckman, P., Sun, D., & Jiang, Y. (2005). A hierarchical model for estimating response time distributions. *Psychonomic Bulletin & Review*, 12, 195-223.
- Schwarz, W. (1989). A new model to explain the redundant-signal effect. *Perception and Psychophysics*, 46(5), 498-500.
- Smith, E. E. (1968). Choice reaction time: An analysis of the major theoretical positions. *Psychological Bulletin*, 69, 77-110.
- Townsend, J. T., & Nozawa, G. (1995). Spatio-temporal Properties of Elementary Perception: An Investigation of Parallel, Serial, and Coactive Theories. *Journal of Mathematical Psychology*, 39, 321-359.

- Treisman, A., & Souther, J. (1985). Search asymmetry: a diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General*, 114(3), 285-310.
- Tsodyks, M. V., & Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proceedings of the National Academy of Sciences*, 94, 719–723.
- Ulrich, R., & Miller, J. (1993). Information processing models generating lognormally distributed reaction times. *Journal of Mathematical Psychology*, 37, 513-525.
- Ungerleider, L. G., & Mishkin, M. (1982). Two Cortical Visual Systems. In D. J. Ingle, M. A. Goodale, & Mansfield, R. J. W. (Eds.), *Analysis of Visual Behavior* (pp. 549-586). Cambridge, MA: MIT Press.
- Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, 108, 550-592.
- Van der Heijden, A. H. C. (1975). Some evidence for a limited capacity parallel self terminating process in simple visual search tasks. *Acta Psychologica*, 39(1), 21-41.
- Van der Heijden, A. H. C., La Heij, W., & Boer, J. P. A. (1983). Parallel Processing of Redundant Targets in Simple Visual Search Tasks. *Psychological Research*, 45, 235-254.

Van der Heijden, A. H. C., Schreuder, R., Maris, L., & Neerincx, M. (1984). Some evidence for correlated separate activation in a simple letter-detection task.

Perception and Psychophysics, 36, 577-585.

Wundt, W. (1880). *Grundzüge der physiologischen Psychologie*. 2nd ed., Leipzig: Engelmann.

Appendix A: Simulating the coactivation model

The coactivation model simulated here accumulated evidence until a common threshold is exceeded. We arbitrarily set the threshold K at 10, i.e. as soon as 10 evidences for a target have entered, the target is recognized. Times to register an evidence were drawn from a Weibull distribution with parameters $\alpha = 0$; $\beta = 10$; $\gamma = 2$. To get a response time for an object with one target attribute, K evidence times were drawn from the evidence distribution, and summed up – the result was the simulated response time for this specific target. This process was repeated a hundred times for that condition to simulate 100 trials.

For targets with redundant attributes, evidences were drawn from the evidence distribution for each channel and summed up as they were drawn (so that the time to register the n th evidence on channel i is the time to register all the previous evidences plus the time drawn from the evidence distribution). Then the K^{th} -smallest time, irrespective of the channel, was selected as the RT to a triple redundant target. Again this process was repeated a hundred time. Hence, coactivation happens because the K^{th} evidence to be accumulated, no matter what channel it arrives from, is sufficient to overcome the threshold for object recognition. The simulated RTs were fitted using a Weibull distribution (using the same procedure as for the empirical data).

The exact parameters we chose for the model are not relevant, since we wanted to investigate trend and not absolute parameter values (e.g. does γ tends to increase, decrease, or stay the same from non-redundant to redundant target

conditions). Further, α and β being scaling parameters, they have no influence on the trends observed. Regarding γ , we repeated the simulations with different values (very skewed evidence time, $\gamma = 1.2$, to near symmetrical evidence time, $\gamma = 3$). There was no qualitative difference in the predicted RT asymmetry as a function of redundancy. We also varied the threshold size K from very small ($K = 3$) to very large ($K = 20$) with no changes in the trends.

Tables

Table 1. Stimulus distribution for one block of 60 trials. Panels, rows and columns show the different attribute values. White fields are non targets, light gray fields are stimuli with one target attribute, medium gray fields are stimuli with two target attributes, and the dark gray field represents a stimulus with all three target attributes present.

Direction of motion											
Right motion				Left motion				Down motion			
Form			Form			Form			Form		
Color	square	circle	triangle	Color	square	circle	triangle	Color	square	circle	triangle
blue	3	3	3	blue	3	3		blue	3		
red	3	3		red	3		3	red		3	3
green	3			green		3	3	green		3	12

Table 2. Proportion of participants for whom the distribution in a condition was significantly different from the distribution in a condition with less redundancy. Between parentheses, the mean Anderson-Darling statistics (AD) for the participants with a significant difference as well as for participants with a non-significant difference.

Double Redundant Conditions	Non-redundant conditions		
	c	f	d
cf	21/24 (11.29/ 1.56) *	17/24 (7.52/ 0.99) *	-
cd	24/24 (19.92/ n.a.) **	-	15/24 (6.36/ 1.41) *
fd	-	23/24 (23.28/ 1.85) *	20/24 (11.36./ 1.33) *

Triple Redundant Condition	Double redundant conditions		
	cf	cd	fd
cfcd	24/24 (22.06/ n.a.)*	18/24 (10.17/ 1.24)*	7/24 (5.90/ 1.20)*

* The critical AD for a decision criterion of .05 is 2.49.

** For one participant, there was a significant difference between the RT distributions for c and cd, but in the opposite direction; i.e. the RT for c was actually faster than for cd. However, this was the only case where such a reversal was observed.

Table 3. Summary of the various predictions made by race models according to their architecture and whether the distribution for a single channel F , is a member of the Power Distribution class or not.

Independent channel model		Prediction		
Architecture	Distribution F for one evidence	Estimated γ across cond.	Estimated α across cond.	Coactivation
k = 1	$F \in \text{PD}$	constant	constant	no
	$F \notin \text{PD}^*$	increasing	decreasing	no
Serial	$F = \text{exponential}$	constant	decreasing	huge
	$F = \text{normal}$	increasing slowly	decreasing	huge
Parallel	$F \in \text{PD}$	almost constant	decreasing	yes
	$F \notin \text{PD}^*$	increasing	decreasing	yes

* F is the lognormal distribution.

Table 4. Fit of the models to the response time distributions of the correct response for each participant across the three redundancy conditions.

Participant	3Lognormal	3Weibull	Cascade race model
11	2559.6	2610.5 *	2575.8 *
12	2475.7	2492.3 *	2481.0
13	2629.6	2638.5	2633.1
14	2589.6	2587.0	2598.2
15	2522.8	2531.3	2535.6 *
16	2532.8	2544.9	2546.3 *
17	2492.4	2520.5 *	2505.8 *
18	2352.9	2378.4 *	2376.0 *
21	2450.8	2478.0 *	2456.9
22	2525.1	2548.4 *	2532.0
23	2480.6	2514.1 *	2506.9 *
24	2487.1	2514.6 *	2492.1
25	2518.3	2545.1 *	2527.5
26	2630.4	2642.6 *	2641.2
27	2487.3	2515.5 *	2505.9 *
28	2509.8	2549.7	2520.1
31	2532.5	2559.0	2543.4
32	2606.7	2620.3 *	2619.4 *
33	2501.0	2533.5 *	2526.0 *
34	2683.0	2690.2	2692.0
35	2518.3	2529.2	2530.7
37	2604.3	2620.3	2616.5
38	2669.0	2675.1	2673.4
39	2407.3	2428.0 *	2415.5 *

* A significantly worse fit than the 3lognormal model using a BIC penalty term

Table 5. Mean fit and mean parameter values for the models tested.

	Mean fit	mean T_0	mean β/μ	mean γ/σ	mean k	mean ρ
3Lognormal	2512.3	185.2	5.115	0.350		
3Weibull	2534.9	278.8	129.0	2.072		
Cascade race model	2523.8	249.6	5.879	0.779	1.298	7.685

Figure Legends

Figure 1. Mean response time per condition over all participants. Error bars show standard deviation.

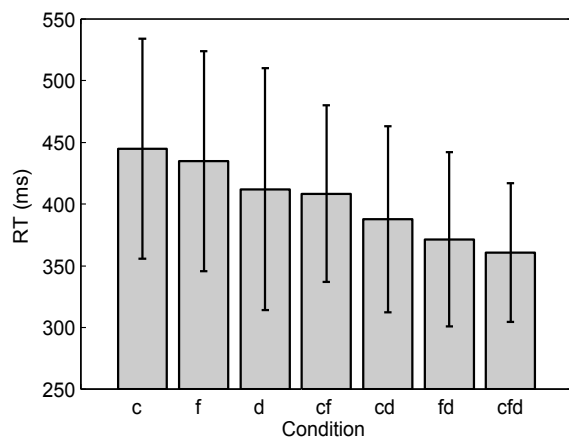


Figure 2. RT distributions of participant 13 for single- and double redundant conditions. Gray lines are the cumulative distributions for single-target conditions, black lines for double redundant conditions. All double redundant conditions for this participant are significantly faster than the non-redundant conditions.

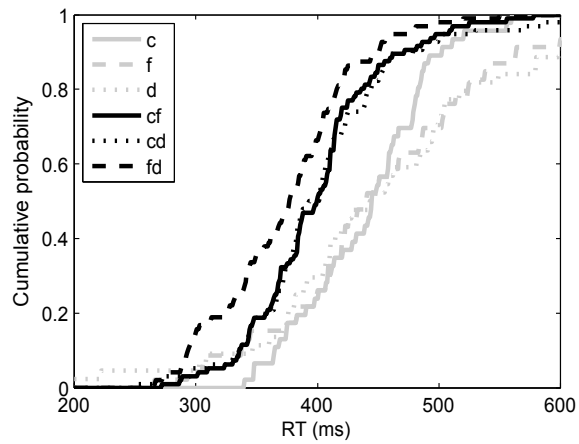


Figure 3. RT distributions of participant 13 for double and triple redundant conditions. Gray lines are the cumulative distributions for double redundant conditions, the black line for the triple redundant condition. This participant is significantly faster in the triple redundant condition than in any of the double redundant conditions.

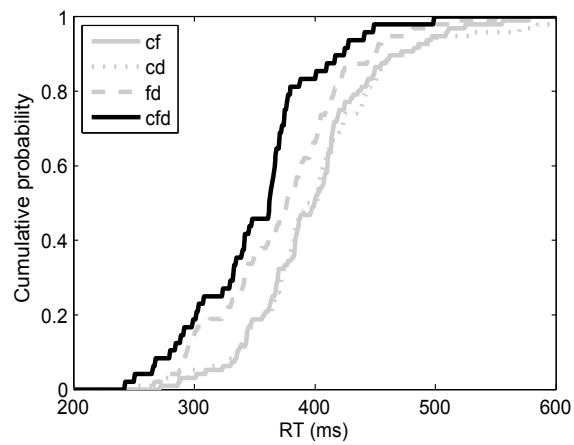


Figure 4. Townsend Bound and RT distributions of participant 13 for double redundant conditions.

Panels represent the three different combinations of two target attributes. Gray lines are the cumulative distributions for single-target conditions, the continued black line the respective double redundant condition and the dotted black line shows the corresponding Townsend Bound. This participant responds significantly faster than the corresponding Townsend Bound only for the condition fd (see panel C). In the other two conditions (Panels A and B) there is no significant difference between the Townsend Bound and actual RT distribution.

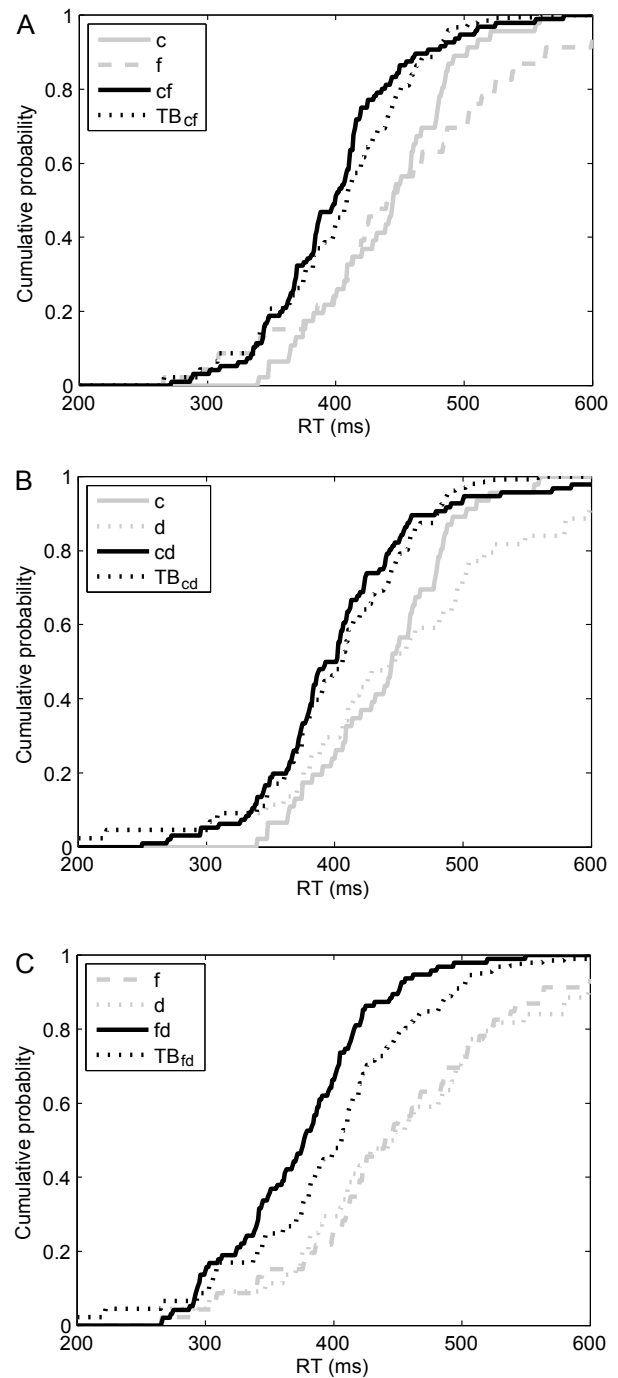


Figure 5. Townsend Bound and RT distributions of participant 13 for triple redundant conditions. Gray lines are the cumulative distributions for single-target conditions, the continued black line the triple redundant condition and the dotted black line shows the corresponding Townsend Bound. The participant is significantly faster in the triple redundant condition than predicted by the corresponding Townsend Bound.

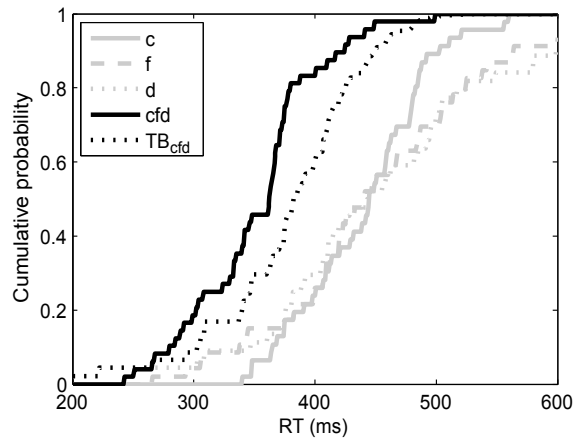


Figure 6. Townsend Bound calculated from double redundant distributions for participants 25 (panel A) and 39 (Panel B). Gray lines are the cumulative distributions for double redundant conditions; the continued black line the triple redundant condition, the dotted black line shows the corresponding Townsend Bound, and the dashed black line the corresponding Townsend Bound calculated from double redundant conditions (TB_{3over2}). Both participants' RTs are significantly faster in the triple redundant condition than predicted by both Townsend Bounds.

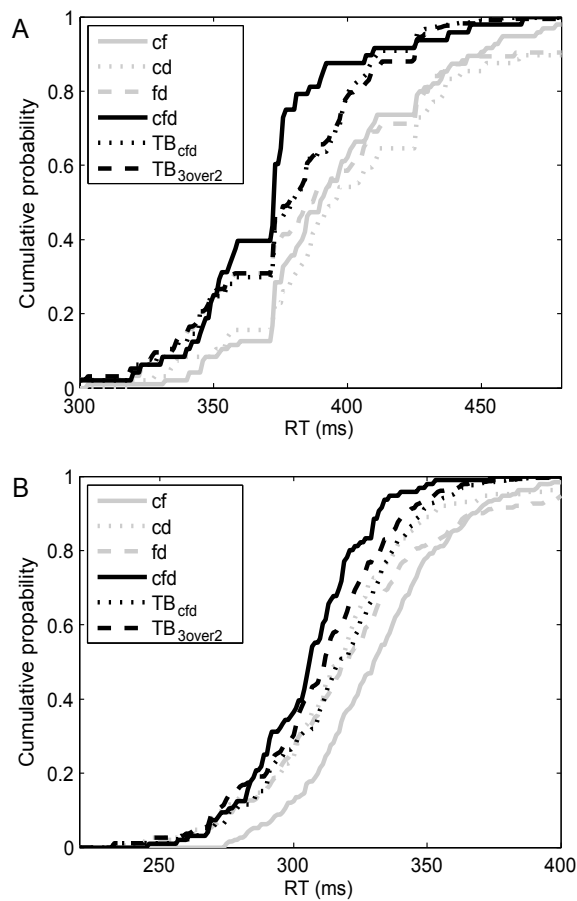


Figure 7. Parameters of RT distributions. Panel A shows the mean over participants of α as a function of redundancy (error bars represent the between-participants standard error; see Loftus & Masson, 1994; Cousineau, 2005), panel B the mean of β , and panel C the mean of γ . For redundant conditions, the darker bar shows participants that were significantly faster than predicted by the Townsend Bound, the lighter bars represent participants that were not.

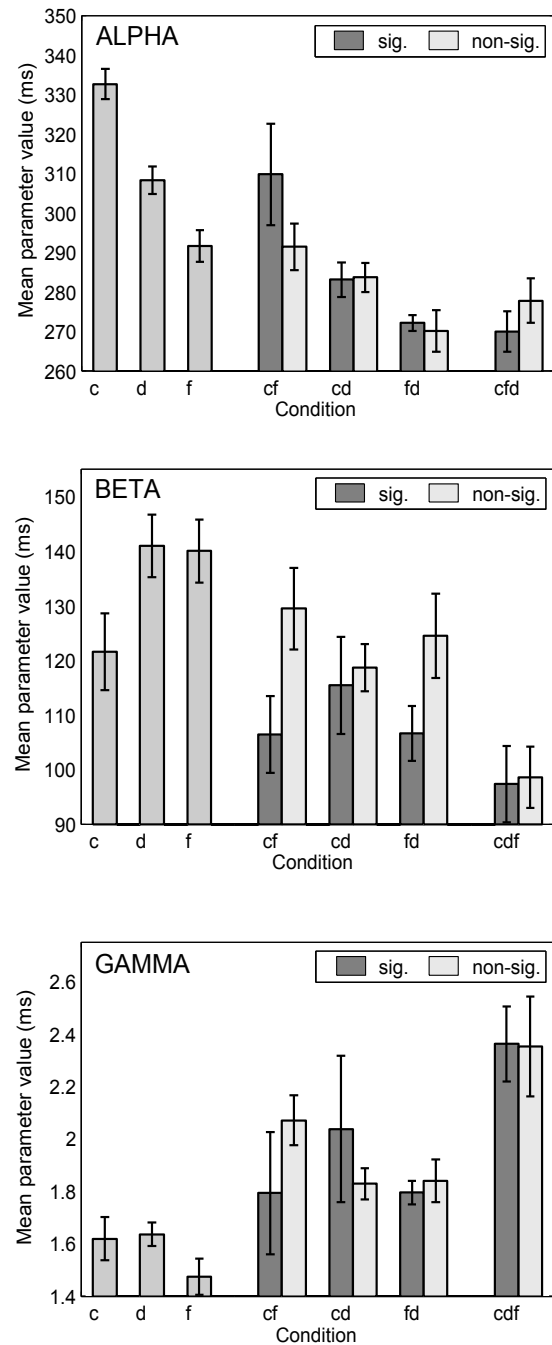


Figure 8. Parameters of RT distributions – real and simulated. Panels A-C show the mean parameter value fitted to participants' RT distributions (error bars represent the between-participants standard error; see Loftus & Masson, 1994; Cousineau, 2005); panels D-F show the mean parameter values of the RT distributions simulated by the coactivation model (error bars represent the standard deviation across simulations). The first column shows the change in α with increasing redundancy, the second column the change in β , and the third column, the change in γ .

(see next page for figure)

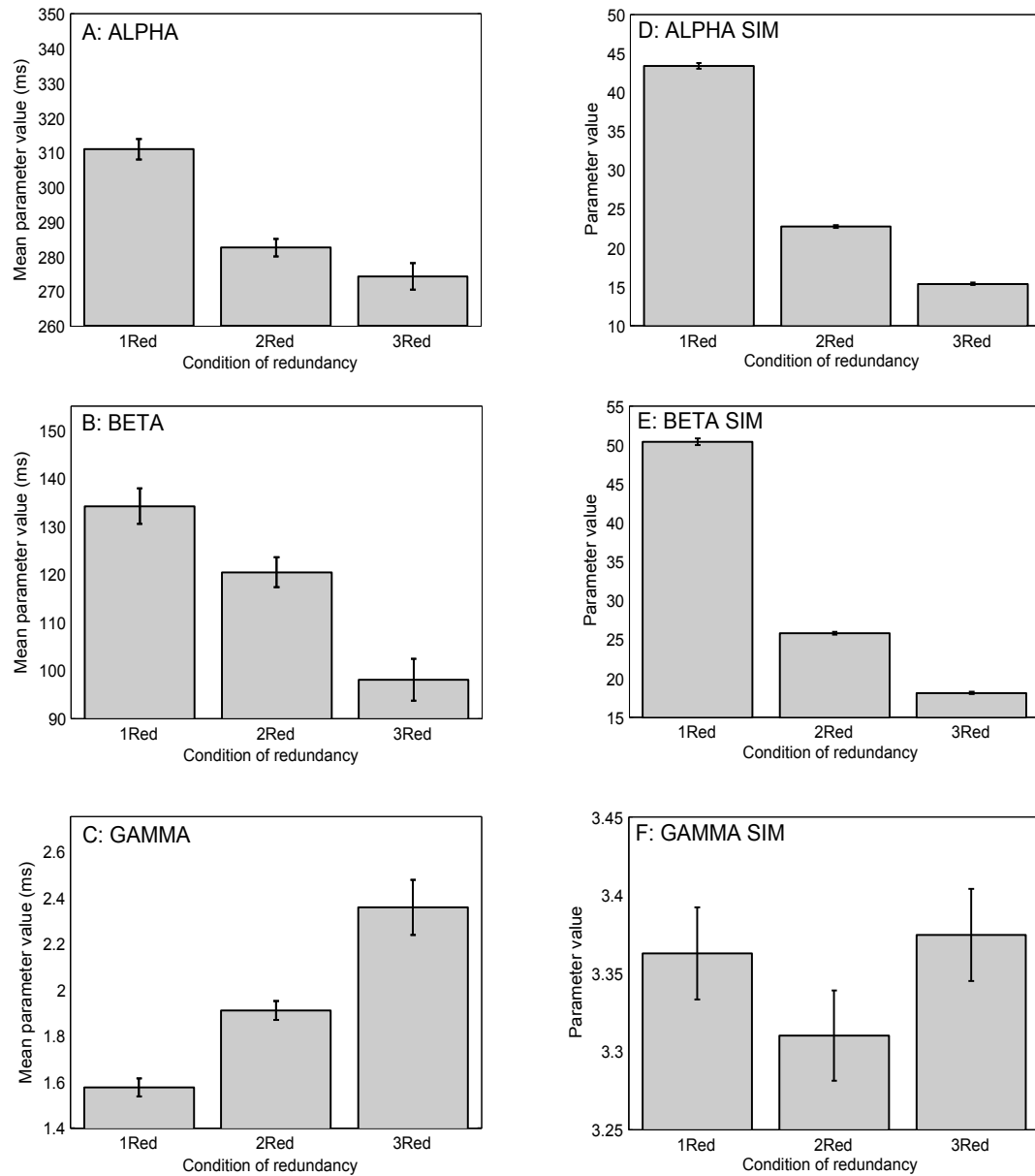
Figure 8. (see previous page for legend)

Figure 9. Different versions of a race model with three possible inputs. Top: a single evidence threshold. Bottom: Race models based on a decision threshold larger than 1. Bottom left: evidences from a cue are received serially, in which case the distribution F characterizes the inter-evidence times; Bottom right: evidences from a cue are received in parallel, in which case the distribution F characterizes a single channel.

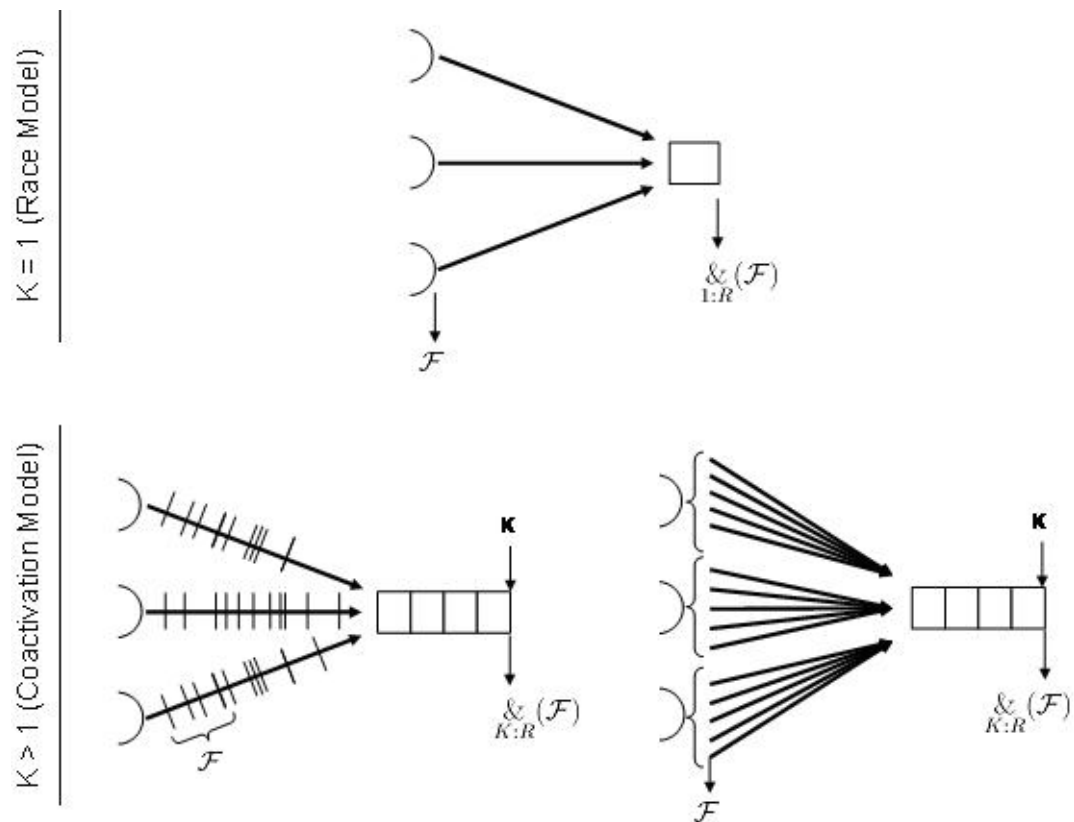
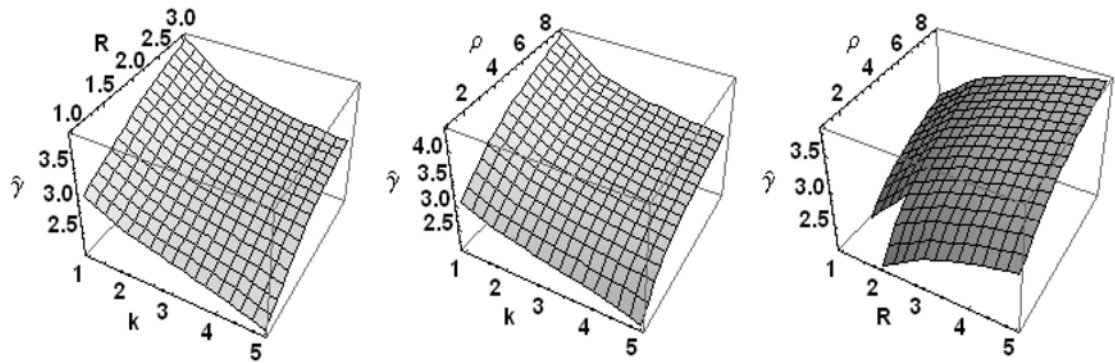


Figure 10. Influence of the parameters ρ (internal redundancy), R (external redundancy) and k (decision threshold) on the observed symmetry as measured by fitting a Weibull to the predicted distribution.



CHAPTER 3

Comparing distributions: the two-sample Anderson-Darling test as an alternative to the Kolmogorov-Smirnoff test

Sonja Engmann and Denis Cousineau

Université de Montréal

Manuscript submitted to the journal *Behavior Research Methods*

Déclaration des coauteurs d'un article

4. Identification de l'étudiante :

Sonja Engmann

Ph. D Psychologie, option Sciences Cognitives et Neuropsychologie

5. Description de l'article :

Engmann, S., & Cousineau, D. (submitted). **Comparing distributions: the two-sample Anderson-Darling test as an alternative to the Kolmogorov-Smirnoff test.** Submitted to *Behavior Research Methods*.

6. Déclaration de tous les coauteurs autres que l'étudiante :

À titre de coauteur de l'article identifié ci-dessus, je suis d'accord pour que Sonja Engmann incluse l'article identifié ci-dessus dans sa thèse de doctorat, qui a pour titre : « Redundancy gain : manifestations, causes and predictions ».

Denis Cousineau
Coauteur

Signature

Date

Apport original

Apport original de Sonja Engmann à l'article :

“Comparing distributions: the two-sample Anderson-Darling test as an alternative to the Kolmogorov-Smirnoff test.”

J'ai conçu et préparé les simulations Monte Carlo présentées dans cet article, ainsi que leurs analyses, sous la supervision et avec les conseils de Denis Cousineau.

L'écriture de l'article a été faite en collaboration avec le Dr Cousineau.

Abstract

This paper introduces the two-sample Anderson-Darling (AD) test of goodness of fit as a tool for comparing distributions, response time distributions in particular. We discuss the problematic of pooling response times across participants, and alternative tests of distributions, the most common being the Kolmogorov-Smirnoff (KS) test. We compare the KS test and the AD test, presenting conclusive evidence that the AD test is more powerful: when comparing two distributions that vary (1) in shift only, (2) in scale only, (3) in symmetry only, or (4) that have the same mean and standard deviation but differ on the tail ends only, the AD test proves to detect differences better than the KS test. In addition, the AD test has a type I error rate corresponding to *alpha* whereas the KS test is overly conservative. Finally, the AD test requires less data than the KS test to reach sufficient statistical power.

Introduction

The motivation for this article lies in the authors' own research on redundancy gain (Miller, 1982): we investigate response time (RT) distributions in an object recognition task, varying the number of redundant attributes identifying an object as a target (Engmann & Cousineau, submitted). We analyze each participant's RTs individually, and therefore need a test that would allow analysis of a whole distribution, not just the mean and variance. We wanted a test that is sensitive to changes in shape and asymmetry. After trying different goodness-of fit tests, we finally settled on the Anderson-Darling test, a powerful tool for comparing data distributions. In this paper we wish to introduce the two-sample version of the Anderson-Darling (AD) test and compare its power to the Kolmogorov-Smirnoff (KS) test.

The AD test is commonly used in engineering, but little known in Cognitive Psychology, despite its advantages for this field. This test is especially useful if there is not a lot of data available in the samples to be compared, and when the analysis should extend beyond distributions' means, taking into account differences in shape and variability as well as the mean of the given distributions. The AD test is non-parametric and can be applied to Normal, Weibull, and other types of distributions. It is especially useful to analyze response time distributions, as it allows a participant-by-participant analysis.

Why combining response time distributions across participants is problematic

When comparing response time (RT) distributions for different experimental conditions, it can be quite difficult to obtain a sufficient amount of data in each condition for a reliable analysis. There is a trade-off between the time participants take for a given experiment and the amount of data per condition. Combining the response times of several participants seems to be, at first glance, an elegant solution to avoid this trade-off. However, on closer inspection, combining RT distributions presents several difficulties.

The most intuitive solution, simply pooling all RTs from all participants together per condition, would produce uninterpretable distributions due to inter-participant variability: such RT distributions would not only be influenced by the characteristics of the experimental condition under which they are produced, but also by individual differences. Participants can have faster or slower motor reactions, or object recognition speed – the possibilities to produce variance in RT distributions are endless – such that variance between participants will be larger than variance due to experimental manipulation. Therefore, simple pooling of different RT distributions will flatten the shape of the final distribution, or, if there are not many participants, lead to a bi- or multimodal distribution.

A technique to avoid some of these problems was proposed by Vincent (1912; see also Rouder & Speckman 2004). The so-called Vincentizing is the most popular technique to combine response time distributions. It involves dividing each distribution into a certain number of quantiles, and then averaging the n th

quantiles of each distribution. The advantage of using this technique is that the resulting “average” RT distribution takes into consideration the relative position of each response time in relation to the other RTs of a specific participant, i.e. minimal RTs are averaged with other minimal RTs; RTs at the peak of each participant’s distribution are averaged with other peaks; etc. This avoids a flattening or multi-modality of the Vincentized distribution.

However, Vincentizing distorts the shape and symmetry of individual distributions (Thomas & Ross, 1980). If an RT distribution reflects one or more underlying processes that contribute to the RT, then this information is essential for analysis. A Vincentized distribution tends towards normality, whereas asymmetry is a universal finding in RT empirical data (Logan, 1992; Rouder, Lu, Speckman, Sun & Jiang, 2005). Possibly relevant information about a RT distribution, such as its degree of symmetry, gets lost when Vincentizing.

Vincentizing is the best technique of combining RT distributions available right now. However, even Vincentizing does not render an unbiased and exact analysis of RT distributions, and research for a better method is in progress, but has not been conclusive so far (Lacouture & Cousineau, in press). Therefore, we need to consider methods available for participant-by-participant analysis.

Different methods of comparing distributions participant-by-participant

The most common methods of comparing two or several distributions, the t-test or the ANOVA, render a judgment of goodness of fit based on the mean and

variance of distributions under comparison. They do not take shape and symmetry into account, which is not specific enough in a lot of cases, for reasons mentioned in the previous section. Also, both tests are parametric, expecting a normal distribution, whereas RTs have a shape close to the Weibull or the Lognormal distribution.

When investigating redundant target recognition RTs, several authors used multiple t-tests on quantiles (Miller, 1982; Mordkoff & Yantis, 1991, 1993, among others). Quantiles (e. g. the 5th percent quantiles) are computed for each participant in the two conditions whose distributions are to be compared. These quantiles are then tested for equality using a t-test. This procedure is replicated for all quantiles at given intervals (e. g. the 10th, the 15th, etc. percent). This method allows an estimate of where RT distributions of all participants differ significantly. It keeps individual participants' data separate, and analyses more than distribution means.

However, sample size for each t-test is only as large as the number of participants in an experiment; therefore statistical power may not be sufficient, especially if the effect size is not very large to begin with. Additionally, between-participant variability might be larger than between-condition differences. Finally, the data at one time point are highly correlated with the data at the previous and following time point, influencing the probability of a type I error rate.

There are several types of non-parametric or distribution-free (they neither depend on the specific form, nor on the value of certain parameters in the population distribution; Massey, 1951) goodness of fit tests that either test if a

sample comes from a given theoretical distribution, or if two samples come from the same underlying distribution. The most well-known in psychology, although used more frequently as a test of independence than goodness of fit, is the Pearson's Chi square (χ^2) test (Chernoff & Lehmann, 1954). The χ^2 test operates on binned frequency distributions, not on probability distributions, and does not give precise results when bin size is too narrow. It is therefore less adapted and less powerful than other tests for comparison of distributions, such as the Kolmogorov-Smirnoff, Cramer-von Mises, Kuiper, Watson or Anderson-Darling test (Stephens, 1974). All of the above tests have more or less the same underlying structure, or are adaptations of one another for different sample sizes or situations, some being more powerful for detecting changes in mean, others in variance (Stephens, 1974).

The Kolmogorov-Smirnoff (KS) test is the most well-known of these tests, and the most commonly used in psychology. The KS test's statistical power is greater than that of the χ^2 -test, it requires less computation, and unlike the latter, it does not lose information by binning, as it treats individual data separately (Massey, 1951; Lilliefors, 1967). However, it is applicable neither for discrete distributions, nor in cases where not all parameters of a theoretical distribution are known and therefore, they have to be estimated from the sample itself.

In this article, we will concentrate on a comparison of the Kolmogorov-Smirnoff (KS) and the Anderson-Darling (AD) test. The former test is already commonly used in the field of psychology, and both are non-parametric,

distribution-free, do not require normality, and are best adapted to the context of RT distribution analysis.

Comparison of Kolmogorov-Smirnoff and Anderson-Darling tests

Both the KS and the AD tests are based on the cumulative probability distribution of data. They are both based on calculating the distance between distributions at each unit of the scale (i.e. time points for RT distributions).

Kolmogorov-Smirnoff Test

The Kolmogorov-Smirnoff (KS) test was first introduced by Kolmogorov (1933, 1941) and Smirnoff (1939) as a test of the distance or deviation of empirical distributions from a postulated theoretical distribution. The KS statistic for a given theoretical cumulative distribution $F(x)$ is

$$KS_n = \sqrt{n} \sup_x |F_n(x) - F(x)| \quad (1)$$

where $F(x)$ is the theoretical cumulative distribution value at x , and $F_n(x)$ is the empirical cumulative distribution value for a sample size of n . The null hypothesis that $F_n(x)$ comes from the underlying distribution $F(x)$ is rejected if KS_n is larger than the critical value KS_α at a given α (for a table of critical values for different sample sizes see Massey, 1951; less conservative critical values exist if the test distribution is the normal distribution, Lilliefors, 1967, or the exponential distribution, Lilliefors, 1969). This means that a band with a height of KS_α is drawn on both sides of the theoretical distribution, and if the empirical distribution falls outside that band at any given point, the null hypothesis is rejected. The KS-

statistic is sometimes abbreviated as D-statistic. For reasons of clarity we will use the former term throughout this article.

The two-sample version of the KS test generalizes to

$$KS_{nn'} = \sqrt{\frac{nn'}{n+n'}} \sup_x |F_n(x) - F_{n'}(x)| \quad (2)$$

where $F_n(x)$ and $F_{n'}(x)$ are two empirical cumulative distribution values at time point x , based on data sets of size n and n' respectively. The null hypothesis that $F_n(x)$ and $F_{n'}(x)$ come from the same underlying distribution is rejected if $KS_{nn'}$ is larger than the critical value KS_α at a given α (for a table of critical values for the two-sample KS test, see Massey, 1951).

The main advantage of the KS test is its sensitivity to the shape of a distribution because it can detect differences everywhere along the scale (Darling, 1957). Also, it is applicable and dependable even for small sample sizes (Lilliefors, 1967). Therefore, a KS test is advised in the following experimental situations: (1) when distribution means or medians are similar but differences in variance or symmetry are suspected; (2) when sample sizes are small; (3) when differences between distributions are suspected to affect only the upper or lower end of distributions; (4) when the shift between two distributions is hypothesized to be small but systematic; or (5) when two samples are of unequal size.

The KS test is fairly well known in the field of psychology, and has been used for a number of different experimental contexts other than a comparison of

response times, such as a comparison of circadian rhythm (Pandit, 2004), an evaluation of exam performance (Rodriguez, Campos-Sepulveda, Vidrio, Contreras & Valenzuela, 2002), or a comparison of economic decision-making (Eckel & Grossman, 1998).

Initially, the authors also used the KS test to compare the response times of participants in an object recognition task where objects could be defined by one, two or three target attributes (Engmann & Cousineau, submitted). However, we began looking for an alternative for the following reasons. First, participants were faster at recognizing objects defined by several target attributes, but the effect was very small. Second, we wanted to compare our data to a model which made certain assumptions about minimal response times, as well as scale and symmetry of response time distributions. We therefore needed a test that would detect small differences at any time point along the distribution, although sample size was not large (48 to 144 per condition). Since we assumed that a substantial part of the effect would show itself in the minimal response times, we needed a test that was especially sensitive to the extrema of a distribution. We finally settled on the AD test as it fulfilled these criteria better than the KS test.

Anderson-Darling Test

The Anderson-Darling test was developed in 1952 by T.W. Anderson and D.A. Darling (Anderson & Darling, 1952) as an alternative to other statistical tests for detecting sample distributions' departure from normality. Just like the KS test, it was originally intended and used mainly for engineering purposes.

The one-sample AD test statistic is non-directional, and is calculated from the following formula:

$$AD = -n - \frac{1}{n} \sum_{i=1}^n (2i - 1)(\ln(x_{(i)}) + \ln(1 - (x_{(n+1-i)}))) \quad (3)$$

where $\{x_{(1)} < \dots < x_{(n)}\}$ is the *ordered* (from smallest to largest element) sample of size n , and $F(x)$ is the underlying theoretical cumulative distribution to which the sample is compared. The null-hypothesis that $\{x_{(1)} < \dots < x_{(n)}\}$ comes from the underlying distribution $F(x)$ is rejected if AD is larger than the critical value AD_{α} at a given α (for a table of critical values for different sample sizes, see D'Agostino & Stephens, 1986).

The two-sample AD test, introduced by Darling (1957) and Pettitt (1976), generalizes to the following formula:

$$AD = \frac{1}{mn} \sum_{i=1}^{n+m} (N_i Z_{(n+m-ni)})^2 \frac{1}{i Z_{(n+m-i)}} \quad (4)$$

where $Z_{(n+m)}$ represents the combined and ordered samples $X_{(n)}$ and $Y_{(m)}$, of size n and m respectively, and N_i represents the number of observations in $X_{(n)}$ that are equal to or smaller than the i^{th} observation in $Z_{(n+m)}$. See Pettitt (1976) for critical values depending on α and sample size. The null hypothesis that samples $X_{(n)}$ and $Y_{(m)}$ come from the same continuous distribution is rejected if AD is larger than the correspondent critical value.

The AD test has been further generalized to a k-sample version (Scholz & Stephens, 1987), which is especially useful to test for the homogeneity of several samples. However, this version will not be discussed in this article.

Several comparisons between the one-sample AD test and other similar tests have been made. Anderson and Darling (1954) found that for one set of observations, the KS and AD test produced the same result. Stephens (1974) compared several one-sample goodness of fit tests, and concluded that while all tests surpassed the χ^2 test in power, the KS, AD, and Cramer-von Mises tests detected changes in mean better.

The AD test has the same advantages mentioned for the KS test in the previous section, namely its sensitivity to shape and scale of a distribution (Anderson & Darling, 1954) and its applicability to small samples (Pettitt, 1976). Specifically, the critical values for the AD test rise asymptotically and converge very quickly towards the asymptote (Anderson & Darling, 1954; Pettitt, 1976; Stephens, 1974).

In addition, the AD test has two extra advantages over the KS test. First, it is especially sensitive towards differences at the tails of distributions (as we will show next). Second, there is evidence that the AD test is better capable of detecting very small differences, even between large sample sizes. This is one of its main advantages in the field of engineering. The goal of the following Monte Carlo simulations is to investigate more rigorously the differences in performance

between the KS test and the AD test, especially concerning small differences between samples and sensitivity to tail differences.

The AD test can be used in the same experimental context as the KS test, but it is not known in the field of psychology, the two-sample version even less than the one-sample version. Rare examples of use of the one-sample AD test in psychology include a test of normality for the distribution of judgments of verticality (Keshner, Dokka & Kenyon, 2006), and a test of normality of platelet serotonin level distributions (Mulder et al., 2004). Apart from our own studies (Engmann & Cousineau, submitted), we are not aware of any further examples of use of the two-sample version.

Comparison of the two tests when shift, scale and symmetry are varied independently

To compare the performance of KS versus AD test, we propose to test if the difference between two sets of data sampled from two minimally different distributions is statistically significant, according to the KS test and according to the AD test. By using theoretical distributions with known parameters, we are able to control the actual size of the difference between the two distributions. This allows us to compare the performance of both tests when distributions are very similar as well as when they are dissimilar. Also, this gives us a tool to observe the effect of change in specific parameters on the performance of both tests. Specifically, we can compare performance when distributions differ only at the extreme ends, but not around the mode, as will be done in the subsequent section.

Method

In a given simulation, we used two populations following Weibull distributions with three parameters,

$$D_1 (\alpha, \beta, \gamma) \tag{5a}$$

$$D_2 (\alpha + \Delta_1, \beta + \Delta_2, \gamma + \Delta_3), \tag{5b}$$

where $\alpha = 200$, $\beta = 80$, and $\gamma = 2.0$. These parameters are typical of speeded response time distributions (Heathcote, Brown & Cousineau, 2004). Δ_1 varied between -60 and 60 , in steps of 4 , Δ_2 varied between -30 and 30 , in steps of 2 , and Δ_3 varied between -1.2 and 1.2 , in steps of 0.08 . In the first simulations, only one parameter varied, whereas the other two remained the same ($\Delta = 0$). For each value of Δ_1 , while maintaining Δ_2 and Δ_3 at 0 , a sample was drawn from D_1 as well as from D_2 . A test of significant difference (with $\alpha = 0.05$) between D_1 and D_2 was then performed, using the KS test and then the AD test. This was repeated $10,000$ times for each value of Δ_1 and subsequently for each value of Δ_2 and Δ_3 as well. For each value of Δ_1 , Δ_2 and Δ_3 we were then able to calculate the probability of finding a significant difference between D_1 and D_2 for the KS test and for the AD test. This procedure was used for sample sizes of 16 , 32 and 64 , typical in experimental psychology.

Insert Figure 1 here

Results

Figure 1 shows the probability for both AD test and KS test of finding a significant difference between D_1 and D_2 when Δ_1 changes, plotted along the abscissa. The three panels represent the different sample sizes. The probability of finding a significant difference is plotted as a function of Δ_1 . If D_1 and D_2 are equal ($\Delta_1 = 0$), the AD test finds a significant difference (type I error) in 4.7% of the cases for sample size $n = 16$, 5.0% for $n = 32$, and 5.2% for $n = 64$. This is approximately the type I error usually allowed for (α). The KS test finds a significant difference in only 1.2% ($n = 16$), 2.2% ($n = 32$), and 3.3% ($n = 64$) of the cases. Hence, the KS test is slightly more conservative, allowing for a smaller proportion of type I errors. On the other hand, when Δ_1 differs from zero, the proportion of type II errors is larger for the KS test, finding no significant difference when distributions are actually different.

Insert Figure 2 here

To illustrate the amount of gain of the AD test over the KS test more clearly, we calculated the difference in probability between the two tests. This was done by subtracting the KS-probability from the AD-probability of finding a

significant difference for each value of Δ_1 . Figure 2 plots the difference as a function of change in Δ_1 , the panels representing sample sizes 16, 32 and 64 respectively. Figure 2 clearly shows that performance of the KS test approaches the performance of the AD test (i.e. the difference approaches zero) only for very large differences between distributions, or when the two distributions are equal (i.e. when $\Delta_1 = 0$). As values of Δ_1 approach intermediate values (near ± 25), there is a systematic and constant gain, sometimes as large as 36% for the AD test over the KS test. The AD test detects as much as a quarter of all differences for certain effect sizes which the KS test could not detect.

Differences in performance between KS test and AD test are more pronounced for small sample sizes. This holds for changes in Δ_1 as well as in Δ_2 and Δ_3 , as will be shown next.

insert Figure 3 here

Figure 3a shows the probability for both AD test and KS test of finding a significant difference between D_1 and D_2 when Δ_2 changes, at a sample size of 64. When D_1 and D_2 were equal ($\Delta_2 = 0$), the proportion of type I errors for the AD test was 0.9% ($n = 16$), 2.0% ($n = 32$), and 3.3% ($n = 64$) respectively. Figure 3b represents the advantage of the AD test over the KS test, again at a sample size of

64. For all sample sizes, the AD test performed as good as or better than the KS test, with a maximal advantage of 4.2% ($n = 16$), 4.9% ($n = 32$) or 4.7% ($n = 64$) respectively.

insert Figure 4 here

Figure 4a shows the probability for both AD test and KS test of finding a significant difference between D_1 and D_2 when Δ_3 changes, at a sample size of 64. The curve is less symmetrical as Δ_3 represents a change in symmetry, and the effect of a negative Δ_3 is not the same as the effect of a positive Δ_3 . When D_1 and D_2 were equal ($\Delta_3 = 0$), the proportion of type I errors for the AD test was 0.7% ($n = 16$), 1.8% ($n = 32$), and 2.9% ($n = 64$) respectively. Figure 4b represents the advantage of the AD test over the KS test, again at a sample size of 64. For all sample sizes, the AD test performed as good as or better than the KS test, with a maximal advantage of 4.6% ($n = 16$), 4.9% ($n = 32$) or 4.6% ($n = 64$) respectively.

When D_1 and D_2 are equal, the KS test has a slightly lower type I error rate, but as soon as samples differ even slightly, the AD test outperforms the KS test for the detection of differences in shift (Δ_1), scale (Δ_2), or symmetry (Δ_3).

Comparison of the two tests when D_1 and D_2 differ in the tails only

As mentioned earlier, one of the strengths of the AD test is its sensitivity to the extreme ends of distributions – the minima and maxima. In order to test its performance specifically at the extrema, we decided to compare distributions that differed only at the extreme ends. The degree of difference between such distributions is extremely difficult to compute, and much less to control. Therefore we selected six instances of two distributions that differ at the extrema, and compared each with a KS and an AD test. One of these distributions was a Weibull, the other a Normal with approximately the same mean and variance as the Weibull. See Table 1 for the exact parameters of each of the six sets of distributions used. Figure 5 shows two such pairs of distributions. Weibulls can be asymmetrical, whereas Normals are symmetrical, which means that an overlap can be obtained for large parts of the distributions, while maintaining a difference at one or both of the extrema.

Insert Figure 5 here

Method

We selected a sample of size 16 from the Weibull and the Normal in each set, tested them for significant difference using the KS and then the AD test. We

repeated this procedure 10 000 times, and then calculated the probability of the AD and the KS test of finding a significant difference. In all other aspects, the procedure is the same as in the previous section.

Insert Table 1 here

Results

The results are shown in Table 1, the last column representing the gain of the AD over the KS test. The AD test is able to detect differences in distributions better than the KS test, even if they are located only at the tail(s) of a distribution.

Sample size needed to reach sufficient statistical power when shift, scale and symmetry are varied independently

Another method to assess the advantage of one statistical method over another is based on statistical power (Cohen, 1992). We will compare the required number of data per cell to reach a target power. Following Cohen (1992), we will use 80% as the target power. The method which requires less data to reach a statistical power of 80% is to be preferred.

We defined the effect size for a shift relative to the standard deviation of the parent distribution. In the following, a small effect size is defined as a change

in the shift (α) of the second distribution by a quantity of 0.25σ and by a quantity of 0.75σ for a large effect size. Table 2 lists the definitions of effect size for the three parameters. Hence, for a Weibull distribution with parameters $\gamma = 2.0$ and $\beta = 80$, the standard deviation is 37 ms and the small effect size is a shift by 9.3 ms ($\alpha \pm 9.3$ ms).

insert Table 2 here

Regarding the scale parameter, there is no convention as to what constitutes a small, medium or large effect size. Hence, we adopted the same effect sizes for changes in scale as for changes in shift. Finally, for the changes in symmetry, a large effect size was defined as a change in the symmetry that would be clearly visible on a plot of the two distributions and a small effect as a change in the symmetry that would be difficult to see. As we saw in the first simulations, power is not symmetrical when the parameter γ is near 2.0. We chose to compare distributions with symmetry parameters of 1.25 and 2.75 ($\gamma \pm 0.75$) for a large difference, 1.50 and 2.50 ($\gamma \pm 0.5$) for a medium difference and finally 1.75 and 2.25 ($\gamma \pm 0.25$) for a small difference. Figure 6 shows the resulting distributions for the two extreme conditions.

insert Figure 6 here

Method

Simulations were run in a fashion similar to the previous ones. We varied the sample size until a power of 80% was reached for each of the two tests, the AD test and the KS test. For most cases, the results are based on 10,000 simulations except when sample size is larger than 100, where the results are based on 25000 simulations so that the results are accurate to the third digit.

Results

The results are presented in Table 3. When the change is in the shift parameter, the net effect is to change the mean of the distribution. Hence, a powerful test should have about the same power as a standard test of means on two groups (e.g. a two-sample t-test). As seen, the number of data needed when the AD test is used (29, 61 and 233 for large, medium and small effect sizes respectively) is the same or slightly smaller than the number of data required by a t-test (29, 64 and 252 for large, medium and small differences in means; Cohen, 1992, Cousineau, 2007). The AD test is more powerful than a t-test when comparing two Weibull distributions; this can be explained by the fact that the left tail of a Weibull distribution is characterized by an abrupt onset. For a small effect size, there is an area of 9.3 ms where there are data in the first sample but none in the second sample. Since the AD test is sensible to differences in tails, it detects this

difference in the left tail efficiently. When the two populations are normal, there is no advantage of the AD test over the t-test. The number of required data is 31, 69 and 272 for large, medium and small effect sizes respectively (based on Monte Carlo simulations with normal distributions).

Table 3 also shows the required number of data when the scale parameter and the symmetry parameter are varied. For changes in shift and scale, the required sample size by a KS test to obtain a statistical power of 80% is close to 50% larger than the sample size when using an AD test. Worst, the KS test is poorest at detecting changes in asymmetry, requiring almost twice as many data than the AD test.

insert Table 3 here

In many psychology experiments, it is not known whether results from two groups produce distributions that differ with respect to their shape, scale, or symmetry, or a combination of the above. Hence, the following could be a reasonable rule of thumb for deciding the sample size to ensure sufficient statistical power: For a given expected effect size, choose the sample size associated with the parameter that requires the largest number of data. For example, if a medium difference is expected between two conditions, not knowing which parameter(s) will reflect the change, a safe approach would be to have 116

data per condition (a change in the scale parameter requires the highest number of data to ensure sufficient power). However, this ideal rule of thumb is limited by practical considerations: Considering that an experimental session generally has no more than 600 trials, that there may be a few erroneous responses that must be removed from the samples, and that there usually are more than two or three different conditions in an experiment, a sample size of 116 per condition might not be practical. If a KS test is used, this number reaches 190, a figure nearly impossible to obtain in any practical experimental design. Note that pooling data between sessions to increase sample size per condition is not recommended unless there are no significant practice effects.

Discussion

In conclusion, we have shown that the AD test is more powerful than the KS test in detecting any kind of difference between samples from two different distributions, all the while maintaining an exact type I error rate of .05. The KS test is overly conservative in comparison. This paper provides three different types of evidence that the performance of the AD test is superior. First, the AD test detects small variations of any one parameter between two distributions more reliably than the KS test. This holds for shift, scale, and symmetry parameters, and for all sample sizes examined. Second, the AD test detects differences at the extreme ends of distributions more reliably than the KS test. Again, this holds even for small sample sizes and when the two distributions largely overlap. Finally, the AD test requires much less data per condition than the KS test in order to obtain sufficient statistical power. Since the AD test further possesses the same advantages as the KS test, and can be applied in the same experimental context, the evidence of its superior performance presented here shows that it should be preferred to the KS test as a tool for comparing distributions.

The AD test is recommended in any experimental context which requires a comparison of samples of continuous distributions, such as response time data, which requires more than a comparison of sample means.

The MatLab (MathWorks, Inc., Natick, MA) version of the two-sample Anderson-Darling test, “adtest2.m” for sample sizes larger than eight for both samples is provided in the Appendix. It requires as input two separate arrays of

data, which do not need to be the same length. Samples are not required to be ordered before serving as input. Optionally, the type I error rate (α) can also be given as the third input. If omitted, the default value is $\alpha = .05$. The output of “adtest2.m” confirms or rejects the null hypothesis that both samples come from the same underlying distribution, supplying the value of the AD statistic and the critical value for the specified α . Please note that the AD test is non-directional, that is it will only give evidence of a significant difference between samples, but not which one of the two is greater or smaller. For details on how to use the one-sample AD test, please refer to Stephens (1974).

Acknowledgments

The authors would like to thank Sophie Callies, Étienne Dusmesnil and Laurence Morissette for their comments on a previous version of the manuscript. This research was supported by the *Conseil pour la recherche en sciences naturelles et en génie du Canada*, the *German Academic Exchange Service (DAAD)* and the *Friedrich-Ebert-Stiftung*.

References

- Anderson, T. W., & Darling, D. A. (1952). Asymptotic theory of certain goodness-of-fit criteria based on stochastic processes. *Annals of Mathematical Statistics*, 23, 193-212.
- Anderson, T. W., & Darling, D. A. (1954). A test of goodness of fit. *Journal of the American Statistical Association*, 49, 765–769.
- Chernoff, H., & Lehmann, E. L. (1954). The use of maximum likelihood estimates in χ^2 tests for goodness-of-fit. *The Annals of Mathematical Statistics*, 25, 579-586.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.
- Cousineau, D. (2007). Implementing and evaluating the nested maximum likelihood estimation technique. *Tutorials in Quantitative Methods for Psychology*, 3, 8-13.
- D'Agostino, R. B., & Stephens, M. A. (1986). *Goodness-of-Fit Techniques*. New York: Marcel Dekker.
- Darling, D. A. (1957). The Kolmogorov-Smirnov, Cramér-von Mises tests. *The Annals of Mathematical Statistics*, 28, 823-838.
- Eckel, C. C., & Grossman, P. J. (1998). Are women less selfish than men? Evidence from dictator experiments. *Economic Journal*, 108, 726–735.

Engmann, S., & Cousineau, D. (submitted). Coactivation results cannot be explained by pure coactivation models.

Heathcote, A., Brown, S., & Cousineau, D. (2004). QMPE: Estimating Lognormal, Wald and Weibull RT distributions with a parameter dependent lower bound. *Behavior Research Methods, Instruments, & Computers*, 36, 277-290.

Keshner, E. A., Dokka, M. S., & Kenyon, R. V. (2006). Influences of the Perception of Self- Motion on Postural Parameters. *Cyberpsychology and Behaviour*, 9(2), 163-166.

Kolmogorov, A. N. (1933). Sulla determinazione empirica di una legge di distribuzione. *Giornale dell'Istituto Italiano degli Attuari*, 4, 83-91.

Kolmogorov, A. N. (1941). Confidence limits for an unknown distribution function. *Annals of Mathematical Statistics*, 12, 461-463.

Lacouture, Y., & Cousineau, D. (in press). How to use MATLAB to fit the ex-Gaussian and other probability functions to a distribution of response times. *Tutorials in Quantitative Methods for Psychology*.

Lilliefors, H. W. (1967). On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, 62, 399-402.

- Lilliefors, H. W. (1969). On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown. *Journal of the American Statistical Association*, 6, 387-389.
- Logan, G. D. (1992). Shapes of reaction-time distributions and shapes of learning curves: a test of the instance theory of automaticity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, 883-914.
- Massey, F. J. (1951). The Kolmogorov-Smirnov test of goodness of fit. *Journal of the American Statistical Association*, 46, 68-78.
- Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, 14, 247-279.
- Mordkoff, J. T., & Yantis, S. (1991). An interactive race model of divided attention. *Journal of Experimental Psychology: Human Perception and Performance*, 17, 520-538.
- Mordkoff, J. T., & Yantis, S. (1993). Dividing attention between color and shape: Evidence of co-activation. *Perception and Psychophysics*, 53, 357-366.
- Mulder, E. J., Anderson, G. M., Kema, I. P., de Bildt, A., van Lang, N. D. J., den Boer, J. A., & Minderaa, R. B. (2004). Platelet Serotonin Levels in Pervasive Developmental Disorders and Mental Retardation: Diagnostic Group Differences, Within-Group Distribution, and Behavioral Correlates.

- Journal of the American Academy of Child and Adolescent Psychiatry*, 43(4), 491-499.
- Pandit, A. (2004). Circadian rhythm variation in attempted suicide by deliberate self-poisoning, and in completed suicide, in central Nepal. *Biological Rhythm Research*, 35(3), 229-233.
- Pettitt, A. N. (1976). A two-sample Anderson-Darling rank Statistic. *Biometrika*, 63, 161-168.
- Rodriguez, R., Campos-Sepulveda, E., Vidrio, H., Contreras, E., & Valenzuela, F. (2002). Evaluating knowledge retention of third year medical students taught with an innovative pharmacology program. *Academic Medicine*, 77(6), 574-577.
- Rouder, J. N., Lu, J., Speckman, P., Sun, D., & Jiang, Y. (2005). A hierarchical model for estimating response time distributions. *Psychonomic Bulletin & Review*, 12, 195-223.
- Rouder, J. N., & Speckman, P. L. (2004). An evaluation of the Vincentizing method of forming group-level response time distributions. *Psychonomic Bulletin and Review*, 11, 419-427.
- Scholz, F. W., & Stephens, M. A. (1987). K-sample Anderson-Darling Tests. *Journal of the American Statistical Association*, 82(399), 918-924.

Smirnoff, H. (1939). Sur les Écarts de la Courbe de la Distribution Empirique.

Receuil Mathématique (Matematicheskii Sbornik), 6, 3-26.

Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons.

Journal of the American Statistical Association, 69, 730-737.

Tatler, B. W., Baddeley, R. J., & Gilchrist, I. D. (2005). Visual correlates of

fixation selection: Effects of scale and time. *Vision Research*, 45, 643-659.

Thomas, E. A. C., & Ross, B. (1980). On appropriate procedures for combining

probability distributions within the same family. *Journal of Mathematical*

Psychology, 21, 136-152.

Vincent, S. B. (1912). The function of vibrissae in the behavior of the white rat.

Behavioral Monographs, 1(5).

Appendix

Implementation of the two-sample Anderson-Darling test in MatLab (MathWorks, Inc., Natick, MA). This implementation assumes sample sizes to be larger than eight. Please refer to D'Agostino and Stephens (1986) for an approximate adjustment of the calculation of the AD statistic for smaller sample sizes, or to Pettitt (1976) for a table of critical values of the AD statistic for smaller sample sizes.

```
function [H, adstat, critvalue] = adtest2(sample1, sample2, alpha)
% ADTEST2: Two-sample Anderson-Darling test of significant difference.
% This test is implemented for sample sizes larger than 8. For smaller
% sample sizes please refer to A.N.Pettitt, 1976 (A two-sample
% Anderson-Darling rank statistic) for the critical values.
%
% CALL:          adtest2 (sample1, sample2);
%   [H,adstat,critvalue] = adtest2 (sample1, sample2, alpha);
%   Sample1 and sample2 are the samples to be compared. They must
%   be vectors of a size greater than 8. Alpha specifies the
%   allowed error. If alpha is not specified, a default value of
%   0.05 for alpha is used. Alpha must be either 0.01, 0.05 or 0.1.
%
% RETURN: H gives the statistical decision. H = 0: samples are not
%         significantly different. H = 1: sample1 and sample2 are
%         significantly different (i.e. do not arise from the same
%         underlying distribution).
%         adstat returns the ADstatistic of the comparison of the two
```

```

%      samples. If adstat is greater than the critical value,
%      the two samples are significantly different.
%      critvalue returns the critical value for the alpha used
%
% (c) Sonja Engmann 2007

if nargin < 2, error('Call adtest2 with at least two input arguments'); end
if nargin < 3, alpha = 0.05; end

% Assignment of critical value depending on alpha
if alpha == 0.01, critvalue = 3.857;
elseif alpha == 0.05, critvalue = 2.492;
elseif alpha == 0.1, critvalue = 1.933;
else error('Alpha must be either 0.01, 0.05 or 0.1.');
```

```

end

samplecomb = sort([sample1 sample2]);
ad = 0;
for i = 1:length(samplecomb)-1
    m = length(find(sample1(:)<=samplecomb(i)));
    ad = ad + (((m*length(samplecomb) - length(sample1)*i)^2)/(i*(length(samplecomb)-i)));
end

adstat = ad/(length(sample1)*length(sample2));

if adstat > critvalue, H = 1; else H = 0; end

```

Tables

Table 1. Parameters of the Weibull and Normal distributions from which samples are drawn for comparison. The last three columns show the probability (over 10 000 instances) of finding a significant difference between samples, either by the AD test or the KS test. The last column represents the advantage of the AD over the KS test.

	Weibull parameters			Normal parameters		Probability of finding a significant difference		
	α	β	γ	μ	σ	AD test	KS test	AD - KS
1	0	10	1.5	6	5.4	.200	.032	.168
2	0	10	2.5	8.5	4	.051	.005	.046
3	0	20	1.3	7.5	11.25	.561	.165	.396
4	0	20	4.0	17.5	6.75	.072	.013	.059
5	0	30	1.6	17.5	15.73	.252	.054	.198
6	0	30	2	22.5	14	.087	.018	.069

Table 2. Definition of large, medium and small effect size for the three parameters of the Weibull distribution.

	Definition		
	Large	Medium	Small
α	0.75σ	0.5σ	0.25σ
β	0.75σ	0.5σ	0.25σ
γ	± 0.75	± 0.50	± 0.25

Table 3. Number of data required to reach a power of 80% as a function of the effect size and the test used.

	<u>The Anderson-Darling test</u>			<u>The Kolmogorov-Smirnoff test</u>		
	Large	Medium	Small	Large	Medium	Small
α	29	61	233	42	92	360
β	58	116	412	81	161	564
γ	48	100	377	83	190	768

Figure Legends

Figure 1. The proportion of significant differences between the two distributions for the AD and KS test as a function of $\Delta 1$ (changes in shift). The horizontal gray line is the boundary of an acceptable type I error rate for a decision criterion of 5%. Panels represent sample sizes 16, 32 and 64 respectively.

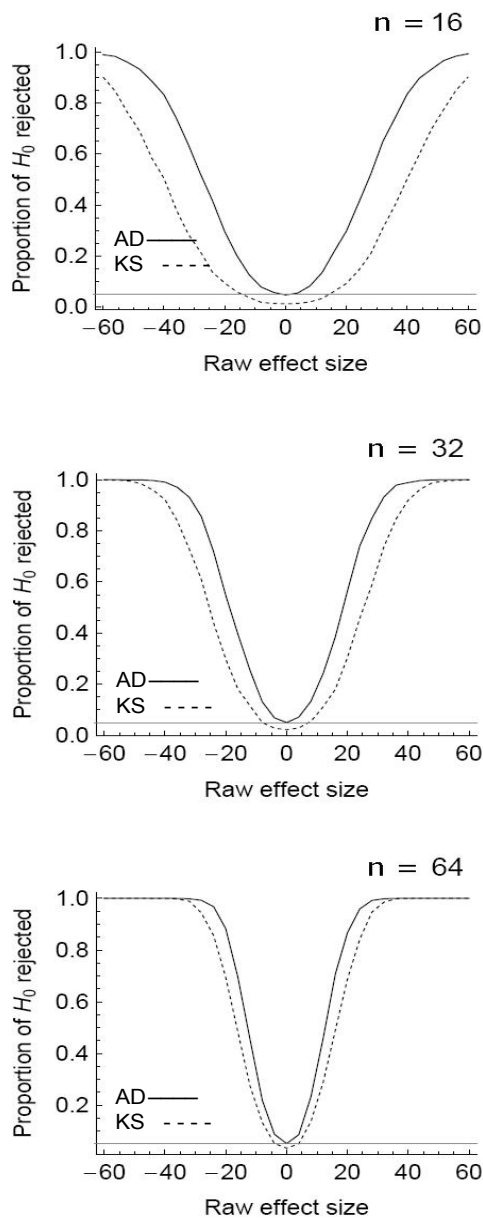


Figure 2. Absolute advantage of AD over KS test as a function of $\Delta 1$ (changes in shift). Panels represent sample sizes 16, 32 and 64 respectively.

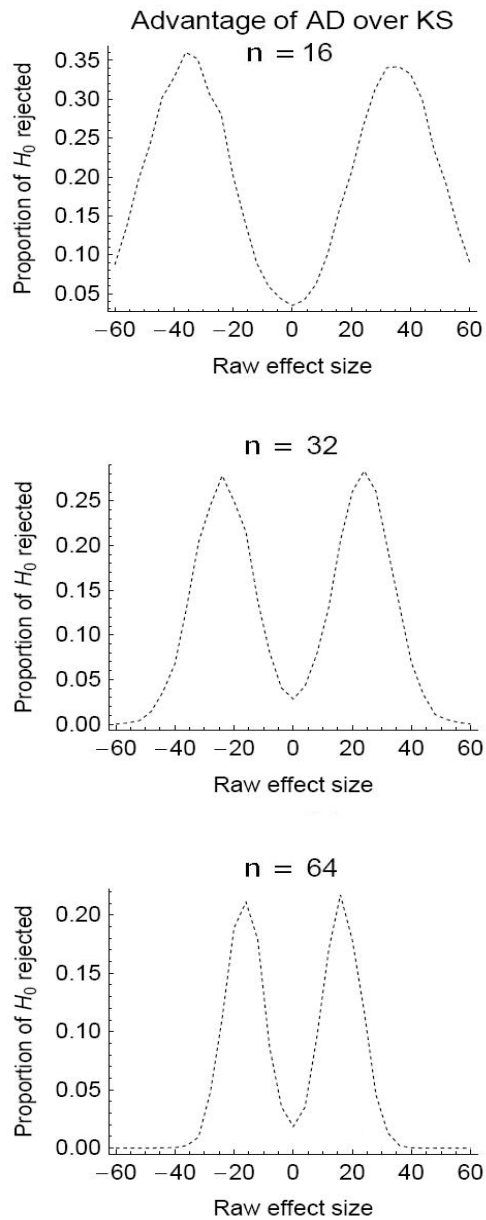


Figure 3. The proportion of significant differences between the two distributions for the AD and KS test as a function of $\Delta 2$ (changes in scale).

The horizontal gray line is the boundary of an acceptable type I error rate for a decision criterion of 5%. The second panel shows the absolute advantage of the AD over the KS test.

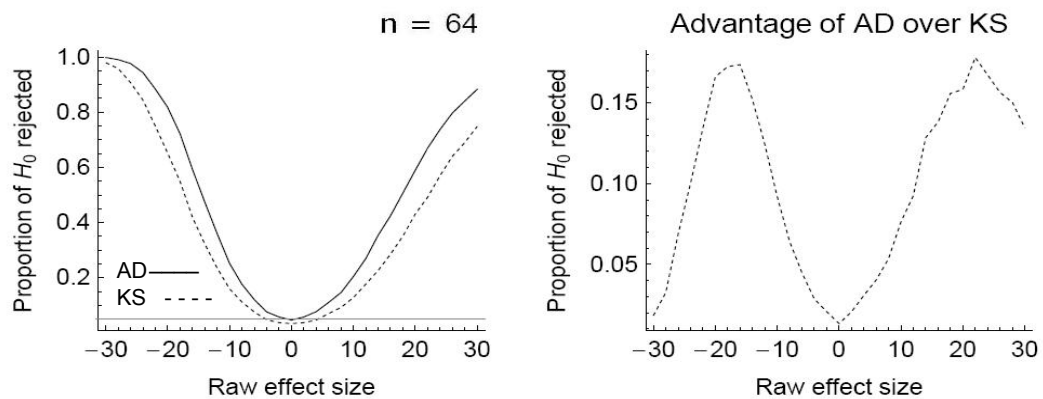


Figure 4: The proportion of significant differences between the two distributions for the AD and KS test as a function of $\Delta 3$ (changes in asymmetry). The horizontal gray line is the boundary of an acceptable type I error rate for a decision criterion of 5%. The second panel shows the absolute advantage of the AD over the KS test.

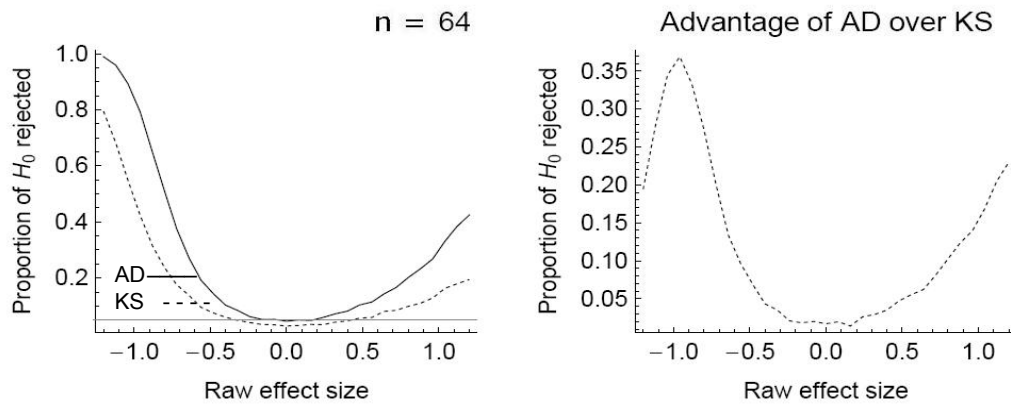


Figure 5. Weibull and Normal distributions used for evaluation of performance when distributions differ at tails. The full line represents the Weibull, the dotted line the corresponding Normal distribution. Panel A shows the pair of distributions for which it was least likely to detect a difference (parameters: Weibull $\alpha=0$, $\beta=10$, $\gamma=2.5$; Normal $\mu=8.5$, $\sigma=4$), panel B the pair for which it was most likely (parameters: Weibull $\alpha=0$, $\beta=20$, $\gamma=1.3$; Normal $\mu=7.5$, $\sigma=11.5$).

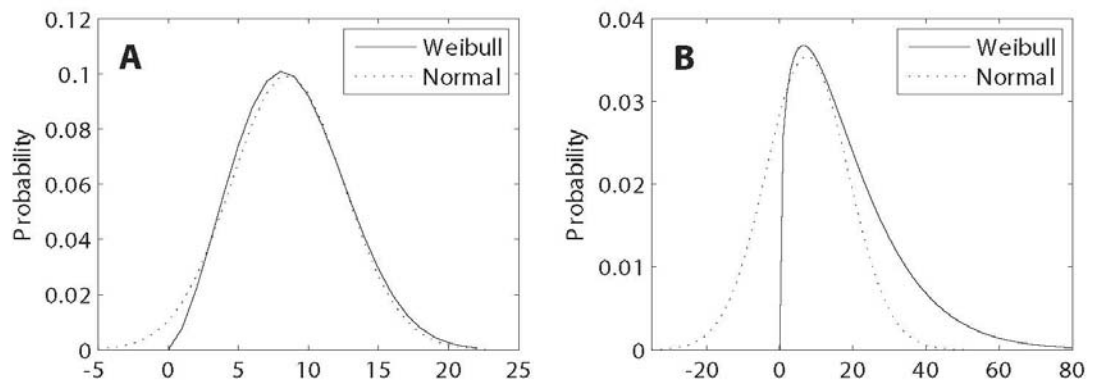
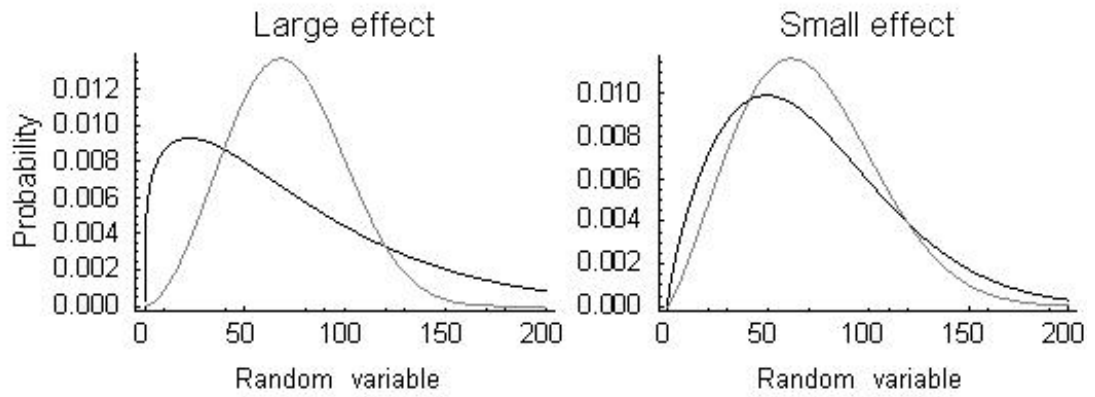


Figure 6. The two distributions compared when the effect size of change in symmetry is large (left) and small (right).



CHAPTER 4

CONCLUSION

The focus of this thesis is a profound comprehension of the characteristics and possible causes of redundancy gain in an object recognition task. In this context, we address several novel issues. First, we found evidence for redundancy gain from three attributes inside a single modality. Second, by extending the Townsend Bound to be based on double-redundant performance, we managed to show that it is possible for the third attribute specifically to contribute more gain than can be predicted by race models. Third, we discovered additional characteristics of the RTE (redundant target effect): redundancy modulates not only mean response times, but also minimal response times and symmetry of response time distributions. We were able to use these additional characteristics in order to exclude pure coactivation models as possible causes of redundancy gain: they do not predict a modulation of symmetry. Fourth, we propose a new model of redundancy gain, the cascade race model, a model that is capable of explaining our results as well as other research on redundancy gain. And finally, we introduce the Anderson-Darling test, and show that it is the most accurate and powerful tool for detecting changes in response time distributions, and therefore most adapted as a statistical test in the present context. At the same time we provide an innovative and very complete way of comparing statistical tests.

This study raises some interesting questions for further research, especially concerning the choice of attributes. We formulated the hypothesis earlier that the combination of colour, form and direction of movement was successful in producing a triple redundancy gain because the three attributes are processed on separate pathways. Verifying this hypothesis is beyond the scope of this study, but could be an interesting direction for future research. A more rigorous and in-depth investigation of the processing of visual features might also be able to establish a rule for success in producing a redundancy gain.

Another important step is to replicate previous research findings on redundancy gain and analyse the modulation of symmetry and onset of response time distributions. A generalisation of the modulation observed here to other feature combinations, and tasks (such as a 2AFC (two alternative forced choice) paradigm), would confirm the validity of an increase in symmetry or a decrease in onset of response time distributions as manifestations of redundancy gain. Additionally, these replications would serve to test the predictions of the cascade race model against larger data sets.

An exploration of the amount of gain added with each new redundant attribute, either with empirical or modelling data, might provide interesting results. Increase in redundancy gain could be linear (each new attribute contributes an equal amount of gain) or asymptotical (redundancy gain reaches a threshold after which no improvement is possible). In light of the present results, the former seems rather unlikely, since performance on single target trials already did not

leave much room for improvement. If the latter is true, the question of where to situate the threshold of redundancy gain remains – are three, four, five or more attributes the limit of information the brain can use to its advantage?

Finally, studying the effects of redundancy on various stages of the processing pathway could find out where redundancy gain is (partially) caused, and help unify research concerning the locus of redundancy gain (see chapter two for a brief review of existing research). This type of research question is however most likely beyond the scope of psychophysics, and more suited for the field of neuroscience. Even though the effects of redundancy on response times are very well documented, we still know almost nothing about the underlying neural structures that cause it.

GENERAL REFERENCES

- Anderson, T. W., & Darling, D. A. (1952). Asymptotic theory of certain goodness-of-fit criteria based on stochastic processes. *Annals of Mathematical Statistics*, 23, 193-212.
- Barlow, H. B. (1972). Single units and cognition: a neuron doctrine for perceptual psychology. *Perception*, 1, 371 -394.
- Biederman, J. (1990). Higher-level vision. In D. N. Osherson, S. M. Kosslyn, & J. M. Hollerbach, (eds.): *An Invitation to Cognitive Science, Volume 2: Visual Cognition and Action*. Cambridge, MA: MIT Press.
- Born, R. T., & Bradley, D. C. (2005). Structure and function of visual area MT. *Annual Review of Neuroscience*, 28, 157–189.
- Cohen, A. L., & Nosofsky, R. M. (2000). An Exemplar-Retrieval Model of Speeded Same-Different Judgments. *Journal of Experimental Psychology: Human Perception and Performance*, 26(5), 1549-1569.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.
- Darling, D. A. (1957). The Kolmogorov-Smirnov, Cramér-von Mises tests. *The Annals of Mathematical Statistics*, 28, 823-838.
- Desimone, R. (1991). Face-selective cells in the temporal cortex of monkeys. *Journal of Cognitive Neuroscience*, 3, 1–8.

- Diederich, A. (1995). Intersensory facilitation of reaction time: Evaluation of counter and diffusion coactivation models. *Journal of Mathematical Psychology* 39, 197-215.
- Diederich, A., & Colonius, H. (2004). Bimodal and trimodal multisensory enhancement of reaction time: Effects of stimulus onset and intensity. *Perception and Psychophysics*, 66, 1388-1404.
- Duffy, C. J., & Wurtz, R. H. (1995). Response of monkey MST neurons to optic flow stimuli with shifted centers of motion. *Journal of Neuroscience*, 15(7), 5192–208.
- Engmann, S., 't Hart, B. M., Sieren, T., Onat, S., König, P., & Einhäuser, W. (in press). Saliency on a natural-scene background: Effects of color- and luminance-contrast add linearly. Accepted for publication in *Attention, Perception and Psychophysics*.
- Estes, W. K. (1986). Array models for category learning. *Cognitive Psychology*, 18(4), 500-49.
- Farah, M. J. (1990). *Visual Agnosia: Disorders of Object Recognition and What They Tell Us About Normal Vision*. Cambridge, MA: MIT Press.
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in neurosciences (Regular edition)*, 15(11), 20-25.

- Graziano, M. S. A., Andersen, R. A., & Snowden, R. J. (1994). Tuning of MST neurons to spiral motions. *Journal of Neuroscience*, 14(1), 54–67.
- Hubel, D. H. (1988) *Eye, brain, and vision*. New York: Scientific American Books.
- Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, 160, 106–154.
- Hubel, D. H., & Wiesel, T. N. (1979). Brain Mechanisms of Vision. *Scientific American*, 241, 150-162.
- Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (2000). *Principles of Neural Science*. New York: McGraw-Hill.
- Kinchla, R. A. (1974). Detecting target elements in multielement arrays: A confusability model. *Perception and Psychophysics*, 15, 149-158.
- Köhler, W. (1947). *Gestalt Psychology*. New York: Liveright.
- Kolmogorov, A. N. (1941). Confidence limits for an unknown distribution function. *Annals of Mathematical Statistics*, 12, 461-463.
- Komatsu, H., & Ideura, Y. (1993). Relationships between color, shape and pattern selectivities of neurons in the inferior temporal cortex of the monkey. *Journal of Neurophysiology*, 70(2), 677–94.

- Krummenacher, J., Müller, H. J., & Heller, D. (2001). Visual search for dimensionally redundant pop-out targets: Evidence for parallel-coactive processing of dimensions. *Perception and Psychophysics*, 63, 901-917.
- Lamme, V. A. F. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *Journal of Neuroscience*, 15, 1605-1615.
- Livingstone, M. S., & Hubel, D. H. (1987b). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience*, 7, 3416-68.
- Livingstone, M. S., & Hubel, D. H. (1988). Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science*, 240, 740-749.
- Logan, G. D. (1992). Shapes of reaction-time distributions and shapes of learning curves: a test of the instance theory of automaticity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, 883-914.
- Logothetis, N. K. & Sheinberg D. L. (1996). Visual Object Recognition. *Annual Review of Neuroscience*, 19, 577-621.
- Lueschow, A., Miller, E. K., & Desimone, R. (1994). Inferior temporal mechanisms for invariant object recognition. *Cerebral Cortex*, 5, 523-531.
- Marr, D. (1976). Early processing of visual information. *Philosophical Transactions of the Royal Society of London, Series B*, 275, 483-524.

- Maunsell, J. H. R., & Newsome, W. T. (1987). Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience*, 10, 363-401.
- Maunsell, J. H. R., & van Essen, D. C. (1983). The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *Journal of Neuroscience*, 3, 2563–2586.
- Miller, J. (1978). Multidimensional same-different judgements: Evidence against independent comparison of dimensions. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 411-422.
- Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, 14, 247-279.
- Milner, P. (1974). A model for visual shape recognition. *Psychological Review*, 81, 521-535.
- Mordkoff, J. T., & Yantis, S. (1991). An interactive race model of divided attention. *Journal of Experimental Psychology: Human Perception and Performance*, 17, 520-538.
- Mordkoff, J. T., & Yantis, S. (1993). Dividing attention between color and shape: Evidence of co-activation. *Perception and Psychophysics*, 53, 357-366.
- Motter, B.C. (1994). Neural correlates of attentive selection for color or luminance in extrastriate area V4. *Journal of Neuroscience*, 14(2), 178-189.

- Newell, A. & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In: J. R. Anderson (ed.), *Cognitive Skills and Their Acquisition*, 1-55. Hillsdale, NJ: Erlbaum.
- Nosofsky, R. M., Clark, S. E., & Shin, H. J. (1989). Rules and exemplars in categorization, identification and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(2), 282-304.
- Palmer, S. E. (1981). The Psychology of Perceptual Organization: A Transformational Approach. In J. Beck, B. Hope & A. Rosenfeld (eds.), *Human and Machine Vision*, 269-339. Orlando, FL: Academic Press, Inc.
- Raab, D. H. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*, 24(5), 574-590.
- Rouder, J. N., Lu, J., Speckman, P., Sun, D., & Jiang, Y. (2005). A hierarchical model for estimating response time distributions. *Psychonomic Bulletin & Review*, 12, 195-223.
- Rouder, J. N., & Speckman, P. L. (2004). An evaluation of the Vincentizing method of forming group-level response time distributions. *Psychonomic Bulletin and Review*, 11, 419-427.
- Roy, J. P., Komatsu, J. & Wurtz, R. H. (1992). Disparity-sensitivity of neurons in extrastriate area MST. *Journal of Neuroscience*, 12, 2478-2492.

- Schwarz, W. (1989). A new model to explain the redundant-signal effect. *Perception and Psychophysics*, 46(5), 498-500.
- Singer, W., & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience*, 1, 555-566.
- Smirnoff, H. (1939). Sur les Écarts de la Courbe de la Distribution Empirique. *Receuil Mathématique (Matematicheskii Sbornik)*, 6, 3-26.
- Smith, E. E. (1968). Choice reaction time: An analysis of the major theoretical positions. *Psychological Bulletin*, 69, 77-110.
- Tarr, M. J., & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology*, 21(28), 233-282.
- Thomas, E. A. C., & Ross, B. (1980). On appropriate procedures for combining probability distributions within the same family. *Journal of Mathematical Psychology*, 21, 136-152.
- Torralba, A., Murphy, K. P., Freeman, W. T., & Rubin, M. A. (2003). Context-based vision system for place and object recognition. *Computer Vision Proceedings. Ninth IEEE International Conference*, 1, 273-280.
- Townsend, J. T., & Nozawa, G. (1995). Spatio-temporal Properties of Elementary Perception: An Investigation of Parallel, Serial, and Coactive Theories. *Journal of Mathematical Psychology*, 39, 321-359.

- Treisman, A. (1990). Features and Objects in Visual Processing. In I. Rock (ed.), *The Perceptual World: Readings from Scientific American*. New York: W. H. Freeman and Co.
- Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology*, 6, 171-178.
- Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, 12, 97-136.
- Treisman, A., & Souther, J. (1985). Search asymmetry: a diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General*, 114(3), 285-310.
- Van der Heijden, A. H. C., La Heij, W., & Boer, J. P. A. (1983). Parallel Processing of Redundant Targets in Simple Visual Search Tasks. *Psychological Research*, 45, 235-254.
- Vincent, S. B. (1912). The function of vibrissae in the behavior of the white rat. *Behavioral Monographs*, 1(5).
- Wang, Q. Q., Cavanagh, P., & Green, M. (1994). Familiarity and pop-out in visual search. *Perception and Psychophysics*, 56, 495-500.

- Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: an alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 419-433.
- Zeki, S. M. (1977). Colour coding in the superior temporal sulcus of rhesus monkey visual cortex. *Proceedings of the Royal Society of London, Series B*, 197, 195-223.
- Zipser, K., Lamme, V. A. F., & Schiller, P. H. (1996). Contextual modulation in primary visual cortex. *Journal of Neuroscience*, 16, 7376–7389.

APPENDIX

Instructions for participants

These instructions were received by participants of the experiment 3RedE prior to the experimental session. Participants received similar instructions prior to all other experiments mentioned in this thesis. Variations included the description of the stimuli and the duration of the experiment.

INSTRUCTIONS AUX PARTICIPANTS

Bienvenu à cette expérience. Elle se tient dans le cadre d'une recherche pour un projet de maîtrise. La recherche est conduite par Sonja Engmann et Marie-Frédérique Beaupré, et elle est supervisée par Dr Denis Cousineau, professeur au département de psychologie de l'Université de Montréal. Ces expérimentateurs seront à votre disposition tout au long de l'expérience en cas de besoin.

Cette expérience est constituée d'une session d'environ 1 heure.

Cette expérience a pour but de mieux comprendre le phénomène de la reconnaissance d'objets. Pour l'accomplissement de cette dernière, on vous demande de suivre quelques règles :

- S'abstenir de l'utilisation de votre cellulaire, etc.
- Garder votre cellulaire, télé-avertisseurs, etc. éteints
- Bouger le moins possible. Ne faire rien d'autre que l'expérience. Il est donc interdit de boire ou de manger.
- Si votre acuité visuelle nécessite le port de lunettes ou de lentilles, il serait préférable de les porter.
- Ne pas vous déplacer avec la chaise puisqu'il est important de vous tenir à peu près à la même distance tout au long de l'expérience.
- Répondre seulement avec la main droite

Pendant l'expérience, des objets animés différents vous seront présentés. Ils sont précédés par un point de fixation, qui est dans la même place où apparaîtra l'objet.

Votre tâche consiste à identifier des objets avec des caractéristiques spécifiques : une certaine couleur ainsi qu'une certaine forme et une certaine direction de mouvement. Vous devrez répondre en appuyant sur la barre d'espacement lorsque l'objet présenté est constitué de un ou plus des traits suivantes : que ce soit de couleur bleu et/ou de forme carré et/ou bouge à droit. Si

aucun de ces trois traits n'est présent, vous devrez vous abstenir d'appuyer sur la barre d'espacement. Votre pointage sera indiqué à la suite de chaque essai dans le simple but de garder votre motivation et votre attention élevées. Votre score final sera dévoilé à la fin de l'expérience. Les points sont distribués de la manière suivante :

- 15 points : vous n'avez pas appuyé sur la barre d'espacement et il ne fallait pas appuyer
- 30 points : vous avez appuyé sur la barre d'espacement et il fallait appuyer
- 50 points : vous avez appuyé sur la barre d'espacement en moins de 300 ms et il fallait appuyer
- -350 points : mauvaise réponse! vous avez appuyé sur la barre d'espacement, mais il ne fallait pas appuyer ou vous n'avez pas appuyé, mais il fallait appuyer

On donne un prix pour le meilleur score à la fin de l'expérience !!

À la fin de l'expérience, vous recevrez une rémunération de 8 \$. Si vous le désirez, à la fin de l'expérience, ce sera un plaisir de répondre à vos questions et de vous donner de plus amples informations sur le sujet. Nous tenons à vous rassurer en vous avisant que vous ne serez pas enregistrés de quelques façons que ce soit durant votre participation. Nos salles ne contiennent aucun microphone, ni aucune caméra. De plus, nous n'évaluerons pas votre quotient intellectuel, nous n'établirons pas votre portrait psychologique et nous ne dresserons pas votre profil de personnalité.

Juste avant de commencer l'expérience, nous vous demandons de bien vouloir signer le formulaire de consentement éclairé, en conformité avec la déontologie sur les recherches en psychologie, si vous acceptez d'être volontaire pour cette expérience. Par contre, même avec votre consentement écrit, vous avez l'opportunité en tout temps de cesser l'expérience. De plus, nous vous demanderons quelques renseignements personnels qui resteront confidentiels. Chaque participant sera identifiable que par un numéro attribué à chacun au tout début de l'expérience. Seul ce numéro sera utilisé pour les résultats, donc votre nom ne figurera dans aucune publication de cette recherche.

Sonja Engmann
Denis Cousineau
Université de Montréal
Département de psychologie
Pavillon Marie-Victorin, local E-404-13