Luminance-defined salience – targets among distractors

Hans-Christoph Nothdurft
Visual Perception Laboratory (VPL) Göttingen, Germany

Salience in vision is achieved from various stimulus properties; luminance differences are the most frequent ones in daily life. This study investigated properties of luminance-defined target salience and explored the rules of how salience changes when the target, its surrounding background, or distractors nearby change their luminance. Two experimental sections of the paper present data when subjects matched targets in different surroundings (luminance range 5.5–68 cd/m^2) for equal salience. In a third, computational section these data are compared with predictions from various algorithms. Some observations can be generalized. (1) Salience computations differ between dark and bright targets, and between targets in different rankings to distractors. For dark maximum targets (the target is the darkest item in the scene), salience computation followed the constant-addition principle, that is, targets appeared equally salient when their luminance difference to distractors was the same; the background luminance was (almost) not important. For bright targets, salience computation often followed the *salmin* algorithm, a modification of the constant-ratio principle (Weber contrast); background luminance settings *were* important. But deviations from these rules were also seen. Particularly good performance with bright targets was sometimes obtained with the *averages* of two algorithms. (2) In cross-polarity salience matches (bright targets compared with dark targets) predictions of equal salience matches were strongly improved when luminance scales were power-transformed. This was not the case for same-polarity matches (e.g., dark targets matched against dark targets). (3) Different aspects of salience must be distinguished, such as *discrimination salience* (which distinguishes the target from distractors) and *item salience* (which distinguishes targets from background). In certain stimulus configurations, item salience can be smaller than discrimination salience and may then affect the perceived salience of a target. (4) In computational analysis, quite a few algorithms could explain certain experimental data and failed with others. There was no single "super" algorithm that could reliably predict *all* salience matches of different target combinations. Major findings are demonstrated and the role of salience variations with retinal eccentricity is discussed. *© Author*

Key Words: salience – luminance contrast – Weber contrast – Michelson contrast – Stevens' brightness law – dark, bright – targets, distractors – popout – maximum-minimum paradigm

Published online: 6-Jun-2015

<u>*Citation:*</u> Nothdurft, H.C. (2015). Luminance-defined salience – targets among distractors. VPL-reports, 2, 1-97. www.vpl-reports.de/2/

Hints for reading: Apologies for such a long paper. In fact, it's three (or more) papers, and different sections could be read separately. But most of their content is related, so splitting had not been useful. The shortest overview is in the Abstract. For more details, I am afraid, you have to go through the text. I tried to write everything clear and did enhance major observations, but still there are many details. Since likely not all readers are interested in these details, I will give hints on what the paragraphs are about and what you may want to read or skip. Just look out for more *Hints for reading* in the text. – It may be helpful if you have already read the paper on equal-salient blob arrays (Nothdurft, 2015).

INTRODUCTION

Hints for reading: The Introduction is short. It reminds you that salience is important, graded, and that exact estimates of salience should be appreciated. Look at Fig.1 if you have not heard of the maximum-minimum paradigm.

An important aspect of vision is the relative salience of objects which may bring some items faster into the focus of attention than others (Nothdurft, 2002, 2006b; Töllner, Zehetleitner, Gramann, & Müller, 2011; Treue, 2003; Turatto & Galfano, 2000; Zehetleitner, Koch, Goschy, & Müller, 2013). This difference can be behaviorally important, in particular when attention is spread over highly competitive tasks (Braun, 1994). To understand visual processing it is therefore necessary not only to distinguish salient from non-salient objects but also to measure and identify the relative strength of salience of different objects in a scene.

This is the second of two recent papers from my lab on salience variations with target luminance. While the previous paper (Nothdurft, 2015) addressed the salience of single or multiple identical targets on various backgrounds, the present paper studies the salience of targets among distractors. The investigated parameter is luminance. Does the salience of a target change when distractors are dimmed or lightened up? And how does it change when only the background is varied?

Some puzzling observations about luminance-defined salience are reported in the literature. For example, a bright square is generally more salient than a gray square on the same dark background. But if the gray square is shown with multiple bright squares, it may become more salient than these (Fig. 1; Nothdurft, 2006a). However, there still remains a salience difference between these targets, which is similarly found with a dark target among less dark distractors and the reversed combination. That difference, here referred to as the maximum-minimum paradigm (Fig. 1), has received considerable interest in vision research but to my knowledge no study has yet come up with an exact description of salience variations in such patterns. The original distinction between maximum and minimum targets came from lesion studies (Schiller & Lee, 1991) which reported that rhesus monkeys could normally detect both types of targets but failed particularly in the minimum target condition after lesions of area V4. Also later studies on healthy human observers reported differences between these targets; targets in the *maximum target* condition (cf. Fig. 1) were detected faster than targets in the *minimum target* condition (Braun, 1994; Nothdurft, 2006a; Zenger-Landolt & Fahle, 2001). Even more, if observers simultaneously had to pay attention to a different task, their performance in detecting minimum targets dropped down but not, or less, their performance in detecting maximum targets (Braun, 1994). These findings together had suggested different ways of visual processing; the maximum target was assumed to be found "pre-attentively" from parallel search across all items whereas the minimum target was thought to be detected from an attentive and serial processing of individual items. Later it was noticed, however, that targets in the minimum condition were generally less salient than targets in the maximum condition (with same but exchanged luminance settings). If the salience of the minimum target was increased so that targets in both conditions were about equal-salient, then also the two search conditions became similar and required the same amount of attention and processing time (Nothdurft, 2006a). Variations of salience with maximum and minimum target conditions are not well understood, and to clarify the basis of this distinction was the initial motivation for the present study.

The more general goal of the study, however, was to establish a systematic description of luminance-defined target salience in multiple-item scenes. This goal turned out to be far more difficult than originally thought and the study did not lead to one simple rule or formula for the computation of luminance-defined salience. That was unexpected for two reasons. First, salience estimates of single or multiple targets without distractors (Nothdurft, 2015) had led to fairly clear and reliable computational rules, even though different such rules were found to be working in different conditions. Second, variations of illumination, and hence variations in target, distractor, and background luminance are common in our environment (sunshine, clouds, dimming lights) and one might have assumed that the visual system would handle such variations in a simple, general way. Instead, the experimental data suggest various rules and algorithms when target salience is computed, depending on the relative brightness of targets and distractors, on their luminance polarity to the background and to some extent also on the intention of the observer. But this does not mean that target salience from luminance were an entirely arbitrary and subjective value. Rather, different observers reported similar perceptual sensations when asked about the relative strength of target salience, and produced similar adjustments when asked to match two targets in salience. Thus, the computation of target salience from target, distractor, and background luminance appears to follow common rules even if these rules could not be generalized for all test conditions.

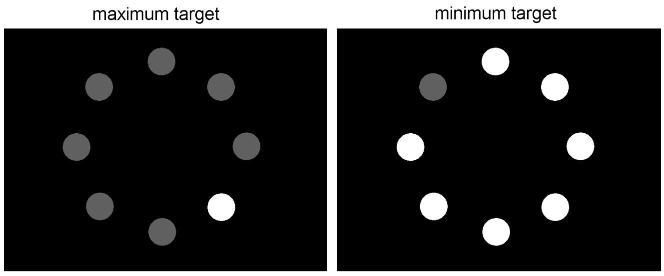


Figure 1. *The maximum-minimum paradigm.* In the maximum configuration on the left, the single bright target is the brightest item in the display; all other items (distractors) are dimmer. In the minimum configuration on the right, the single target is less bright than distractors. When luminance settings of targets and distractors are exchanged, as in the examples shown here, the maximum target is more salient than the minimum target, although the luminance contrast between targets and distractors is the same. The same is true for dark items on brighter background, not shown here. One goal of the study was to measure and predict target salience as a function of luminance contrast to distractors and background.

Experiments were performed on artificial patterns showing a single target at one luminance, and several identical distractors at another luminance level; all items were presented on a background at a third luminance level. Salience comparisons were made between two such patterns by asking subjects to adjust the salience of the target in one pattern to that of the target in the other pattern.

METHODS

Hints for reading: The Method section is important if you want to understand what was done. Read the overview and the section Computational Predictions. Look at Fig.2.

General Overview

The major test procedures were similar to those in the previous paper (Nothdurft, 2015) except that the homogeneous blobs were now distinguished in one target and several distractors. Subjects saw two blob patterns side-by-side and adjusted the target in one pattern (the *test pattern*) to look equal-salient to the target in the second pattern (the *reference pattern*). Subjects were instructed to concentrate only on the salience and conspicuity of items, not explicitly on their lightness, brightness, luminance, or contrast. The results were later compared with predictions from hypothetical salience algorithms. The deviation of predicted from measured data was taken as an indicator of the fitting quality of each computational algorithm.

Stimuli

Stimulus patterns were made of regular blob arrays (Fig. 2); the central blob was the target. Targets and distractors were similar in form (squares, 0.4 deg by 0.4 deg) but differed in luminance; all distractors were identical. Two raster widths were used in different test series, a *wide* raster of 5 by 7 items at a raster width of 2.1 deg, and a *dense* raster of 9 by 17 items at a raster width of 0.5 deg). Reference and test patterns always displayed the same blob raster, and that raster was not changed within a test series. In one experiment, also a "patched" configuration was used in which targets and distractors attached each other so that the background was not seen; these stimuli showed, in fact, single items on the distractor background.

The two stimulus patterns were presented side by side at a target-to-target distance of 10.4 deg, and were inspected through a gray hard-paper mask with two vertical, rectangular holes (8.9 deg x 14.6 deg) in front of the monitor (cf. Nothdurft, 2015, Fig. 1). Reference and test patterns were pseudo-randomly assigned to either side of the screen; each configuration was equally often presented. Test patterns were identified by a small (0.1 deg diameter) green marker underneath to guide subjects which target they could adjust.

All patterns were computer-generated displays (60 frames/s) presented on a 17-inch monitor 75 cm in front of the observer.

Luminance Measurements. Experiments were performed in a dim-lightened room (wall luminance about 3 cd/m^2) on a monitor with carefully controlled luminance settings. The luminance range of the monitor was adjusted to 5.5–68 cd/m^2.

The luminance settings of each stimulus pattern are represented by three values: background, distractor, and target luminance. Dark targets and distractors were produced by higher, bright targets and distractors by lower background settings in the stimulus pattern. All luminance settings were computer controlled via lookup-tables; the exact luminance value of each setting was measured offline and repeatedly checked during the course of the study. All stimuli were achromatic.

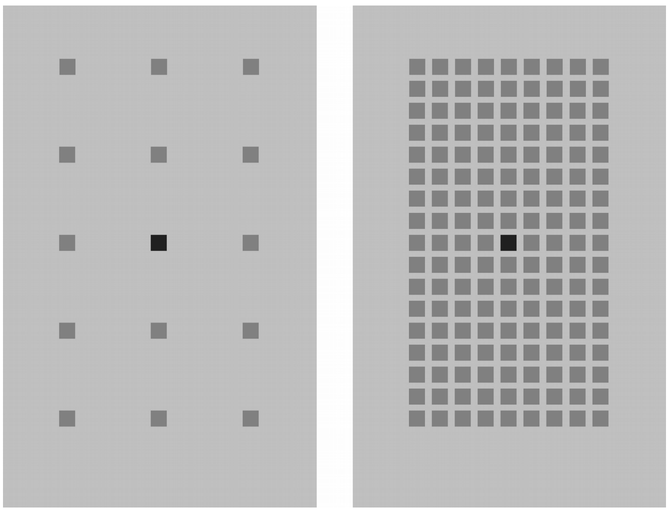


Figure 2. *Examples of stimulus patterns.* Patterns show a DARK maximum target in wide and dense blob arrangements. In the experiments, two such patterns with different luminance settings (but always *same* blob density) were shown side-by-side and had to be adjusted so that both targets appeared equal-salient to the observer.

Procedures

The principle task in all experiments was to adjust the luminance of the test target until it appeared equally salient in its surrounding as the target of the reference pattern, which was fixed and could not be changed. Subjects were encouraged to vary test target luminance over a wide range to explore that either the test or the reference target was the more salient one, before making their final adjustment. The available luminance range was restricted by monitor limitations and was software-blocked at luminance settings where further increments or decrements had changed the target's luminance polarity to background or distractors. For example, subjects could not make a dark target brighter, nor a bright target darker than background and could not convert, e.g., a maximum target into a minimum target.

Subjects made adjustments by pressing the "+" or "–" keys on a computer keyboard, which increased or decreased the luminance difference between test target and test distractors (only targets could be changed); a third key ("a" for "accept") was used to terminate the current adjustment. Upon this response the screen was blanked, and 1s later a new pair of stimulus patterns was shown. If a perfect salience match could not be achieved, subjects were asked to enter the best possible adjustment. During the matches, subjects were asked to shift their gaze between the targets. All stimulus patterns were shown continuously until the subject terminated the trial, and there was no time pressure to finish adjustments within a given time.

Test series

For every task, a large number of individual test series had been designed, as will be described below. Each test series included different stimulus combinations, which were (in random sequence) repeatedly presented during a run. Most test series were repeated once or twice to reduce the standard error of the mean (s.e.m.).

The entire testing program was performed, in an intermingled sequence, on both the *dense* and the *wide* blob raster configurations. Some test series were tested in only one or the other configuration, as indicated below. All measurements were made in two-hour sessions, during which subjects could pause whenever they wanted.

Subjects

Altogether six observers (3 male, 3 female) served as subjects in the various experiments of the study. The main experiments were performed by two (female) students, both 22 years at the beginning of the project, and the author, then 56 years. Later experiments were run by the remaining subjects (27-34 years) and the author. All subjects had normal or corrected-to-normal visual acuity and, except for the author, were naive as to the aim of the study.

Data Analysis

After each trial, the final computer values of background, distractors and target settings were stored and, for analysis, transformed into luminance values taking into account small stray-light effects from medium or high background luminance, as measured on the screen (<2%; see Nothdurft, 2015). Each stimulus pattern was represented by three luminance measures, background luminance, *bg*, distractor luminance, *dis*, and target luminance, *tg*. All computations below are based on these data triplets, one for the reference pattern and one for the test pattern.

Computational Predictions

In order to evaluate the validity of certain algorithms for salience computation, experimental data were compared with a number of predictions from various presumed salience algorithms (Table 1[1]).

For example, if target salience, *sal,* were simply based on the Michelson contrast of targets and distractors, we postulate

(Michelson) $$sal \sim \frac{|tg - dis|}{tg + dis}\text{, or}\quad sal = k \cdot \frac{|tg - dis|}{tg + dis}\ ,$$

with a scaling factor, *k*. For equal salience matches,

(1) $$sal_{reference} = sal_{test}\ ,$$

we can then write

(1a) $$k \cdot \frac{|tg_{reference} - dis_{reference}|}{tg_{reference} + dis_{reference}} = k \cdot \frac{|tg_{test} - dis_{test}|}{tg_{test} + dis_{test}}\ ,$$

1 All Tables are given at the end of the paper to facilitate access throughout the text.

which can be solved for tg_{test} without needing to specify *k,*

$$(2) \qquad tg_{test} = \frac{dis_{test}}{dis_{reference}} \cdot tg_{reference} \ .$$

From the preset luminance values in each stimulus pattern we can thus predict the luminance of an equal-salient test target if salience were related to the Michelson contrast. Note, however, that equation (2) is not necessarily the solution of only one algorithm. For example, the assumption that salience were instead encoded by the Weber contrast of targets and distractors,

$$\text{(Weber)} \quad sal \sim \frac{|tg - dis|}{dis} \ \text{, or} \quad sal = k \cdot \frac{|tg - dis|}{dis} \ ,$$

leads (for some patterns) to the same prediction *(2)*. Conclusions from these predictions must thus be drawn with care.

For each measurement, presumed test target settings were computed for a large variety of algorithms. These predictions were used in two ways. First, for each test series the mean squared deviation (MSD) of predicted from measured data was computed; small MSDs would indicate a good fit of the predicted to the measured data in a series. Second, predicted target settings were plotted and compared with the measured data; the quality of these fits and, in particular, of local deviations were visually inspected.

Note that the MSD values used here have the physical unit $(cd/m^2)^2 = cd^2/m^4$, since they are computed as the squared mean deviation of luminance data (measured in cd/m^2). However, to keep the nomenclature and discussion short and readable, this unit will frequently be left out when *MSD values* are compared.

RESULTS

General Overview

Hints for reading: Read it. This overview is short – and tells you what the main sections are about.

Observations are presented in three parts. *Section A* gives an overview of the experimental results obtained with the equal-salience matching task, when *similar targets* were compared, that is, when both targets were either BRIGHT or DARK and both were presented in either maximum or minimum target configuration. These data are compared with predictions from simple models. – *Section B* expands these experiments to matches of *different items*, like matches of minimum and maximum targets or matches of BRIGHT versus DARK targets. In some of these tests, matches were performed with a fixed test pattern setting, which served as a constant meter and thus allowed to measure and compare the relative salience of various target conditions. – The failure to find a simple unique model that could explain all experimental data led to proof, in *section C*, a much larger variety of algorithms that might underlie salience computation. Predictions from these models are compared with the data. – Discussions in these sections address various aspects of data acquisition, of salience matches, and of conclusions to be drawn from these experiments. Each section is closed with a summary of the main findings and with conclusions from this section. Some sections, and in particular the General Discussion at the end, provide demonstrations of the observed effects, which however need an appropriate printer to be seen (see Appendix).

Note that although the main sections are distinct and separated, there are many links between. Thus, a thorough computational analysis of the data is given in section C, but a first analysis of a few algorithms needs already to be made when the experimental data are presented, that is, in sections A and B. Frequent references are made to Table 1, which lists all major algorithms tested in the study, and to Tables 2-4, which list the best algorithms (smallest MSD values) for the various test series.

A. SALIENCE MATCHES OF SIMILAR TARGETS

Hints for reading: This section reports experimental data, when similar targets (dark or bright) in different luminance settings (background, distractors) are matched for salience. Experiments 1-4 report the general effects of luminance variations of various parameters. Experiments 5 and 6 address details of those matches in higher resolution; you may skip these for an overview.

In all experiments, subjects saw two patterns side-by-side and had to adjust the target in one pattern (the “test” pattern) so that it appeared equally salient to the target in the other pattern (the “reference” pattern). The various test series, which for the clarity of presentation are ordered and sorted into different “Experiments”, were in fact intermixed and tested in interleaved sequences, not in the sequential order they are presented here. Experiments were generally performed by three subjects.

Experiment 1:
Constant distractor luminance on different backgrounds *(Test series block K)*

Stimuli

The test series block *K* included 13 different test series each with 7–12 test conditions. Seven of these series were performed by all three subjects on both wide and dense blob configurations. The remaining test series were performed either by fewer subjects or only on one raster configuration. Within each test series, background and distractor luminance was held constant, and only target luminance was varied between the trials (cf. Fig. 3). Across test series, however, there were systematic variations of background luminance between reference and test patterns. Distractor luminance was always the same in the two patterns (with small deviations from stray light when backgrounds differed strongly, as shown in the plots).

Results and Discussion

Matching results from test series block *K* are shown in Figure 3. Corresponding luminance settings in the reference patterns (black symbols) and test patterns (colored symbols) are plotted side-by-side and in such a way that luminance variations of the reference target fall upon straight lines; the same position shifts are used for plotting test target luminance. Matches from the three subjects were very similar, therefore, only the means and standard errors of the means, *s.e.m.*, are shown. The blue and red data points refer to matches obtained with the wide and the dense blob raster, respectively. Series *K11*, *K22*, and *K43* displayed dark stimuli (backgrounds brighter than targets and distractors), series *K25*, *K26*, and *K47* bright stimuli (backgrounds darker than targets and distractors). Most test series included targets in both maximum and minimum configurations; targets with luminance settings between background and distractors are minimum targets.

Figure 3 illustrates several major findings of Experiment 1. One is that even for continuous luminance variations of the reference target (indicated by the superimposed gray lines) luminance of the salience-matched test targets might have varied discontinuously. That is, the long straight (gray) lines through reference target settings are not accompanied by similar long straight lines through the test target settings, but the latter lines may instead be bent or curved. With *DARK targets* (Fig. 3a), for example, the slopes of luminance variation of reference targets were reproduced in the lower curve sections (*K11* and *K22)* but differed in the upper curve sections (*K11, K22,* and entire test series *K43)* where slopes looked as being scaled to the span of background and distractor luminance (thin black lines). The same response characteristics were seen in other test series of block *K*.

Note that the only difference between reference and test patterns in series block *K* was the different backgrounds. From a first look at these data we might thus hypothesize that the salience of DARK targets is not affected by background luminance when targets are darker than distractors (maximum targets), but is scaled to the luminance span of background and distractors when targets become brighter than distractors (minimum targets). Since this latter case of normalization was, in the present study, initially seen for targets in minimum target configurations, it was named the *"salmin"* prediction (cf. algorithm 5 in Table 1). Only with the low distractor luminance in *K43* were there notable deviations from that prediction.

The situation was different for *BRIGHT targets* (Fig. 3b). Here, the slopes of reference targets were generally not reproduced. Instead, slopes above and below the distractor level now look similar, although different for different backgrounds and distractors. The first impression is that slopes are roughly scaled to the span of background and distractor luminance, as in the *salmin* prediction (thin black lines), but now for targets in minimum and maximum configurations (targets below and above the distractor settings, respectively). Note however, that there are pronounced deviations from the *salmin* predictions in some cases (e.g., *K26, K47)*. In particular, the matches of maximum targets fall *between* the predictions from constant-addition (gray) and the *salmin* algorithm (black). Also the matches of minimum targets in these two series strongly deviate from any of the straight-line predictions.

A second major observation in Figure 3 (beside the just mentioned deviations in test series *K26* and *K47)* is that luminance settings in equal salience matches did not

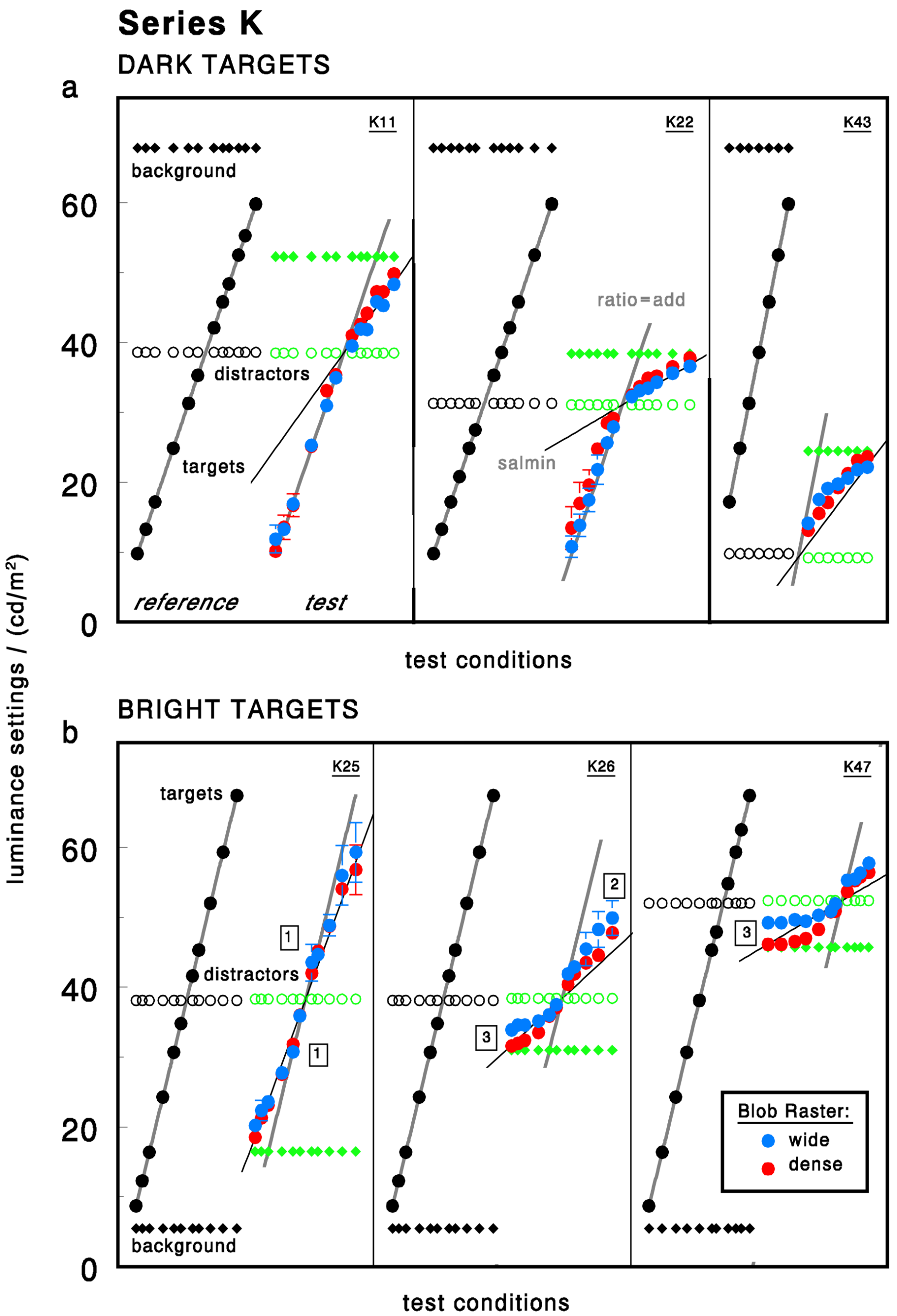


Figure 3. *Luminance settings of salience-matched target conditions in Experiment 1 (test series block K).* Subjects saw two patterns (*reference* and *test)* and adjusted the test target so that it was equally salient to the reference target. Symbols indicate the luminance settings of background (♦), distractors (○), and salience-matched targets (●); reference pattern settings are indicated by black, test pattern settings by colored symbols. Curves show six examples of test series in Experiment 1 *(K11-K47)*; the various reference and test pattern conditions are plotted side by side. **a.** Test series with DARK targets and distractors on brighter backgrounds. **b.** Test series with BRIGHT targets and distractors on darker backgrounds. In each series, data points are shifted along the abscissa so that luminance variations of the reference targets fall upon straight lines; the same shifts are used in the plots of corresponding test pattern data. Graphs show means and s.e.m. (if larger than symbols) of three subjects. Background and distractor luminance settings were constant over each series and only target luminance was varied. In test series block *K,* reference and test distractors were identical; small differences (e.g. in *K43)* are due to stray-light from the different backgrounds. Experiments were performed on wide and dense blob arrangements (blue and red data points, respectively). Straight gray and black lines for test patterns indicate various predictions of the experimental data; for reference patterns they simply confirm the horizontal arrangement to form straight lines. Predictions were based on the constant-addition principle *("add")*, which in these test series was identical with predictions based on the constant-ratio principle *("ratio")*, and on the *"salmin"* algorithm (cf. Table 1). Note that all but one test series include matches of targets in *maximum and minimum configuration* (data points above and below distractor level). For DARK targets, the matches in these conditions follow different slopes; for BRIGHT targets*,* slopes are more similar. Numbered labels mark deviations which were further studied in other experiments.

strongly vary with item density but were similar for the two raster widths tested (red and blue data points). Clearly different matches were obtained with BRIGHT targets in the minimum target configuration when the luminance span of background and distractors was relatively small (as in test series *K26* and *K47*).

Underlying algorithms

To evaluate algorithms that could explain these data, the matching results were compared with predictions from various assumptive salience computations (cf. Table 1). While a thorough analysis of the various algorithms will be made in section C, some computations should already here be referred to, mainly to describe the general matching performances and to illustrate the best fits of the data. To illustrate the interdependence of target variations, I choose the scatter plot presentation in Figure 4.

The similar slopes of salience-matched *DARK targets in maximum condition* (lower curve sections in Fig. 3a) suggest $tg_{test}=tg_{reference}$, which is the prediction based on algorithm 1 and, under the special conditions of Experiment 1 ($dis_{test}=dis_{reference}$), also of algorithms 2 (Weber Contrast) and 3 (Michelson Contrast); see equation *(2)*. To distinguish between these algorithms further variations of stimulus conditions are needed. The fits by these algorithms are rather good and fall, for the curve sections representing maximum targets, upon the dotted identity lines in Figure 4a.

For *DARK targets in minimum condition*, the *salmin* algorithm 5 (thin black lines in Fig. 3a) makes predictions close to the data (except for test series *K43)* but similarly good and partly even better predictions are obtained from other algorithms of Table 1. Figure 4a shows two fits (gray lines) for the DARK minimum targets in test series block *K*; predictions from the *salmin* algorithm 5 (left-hand graph in Fig. 4a) and slightly better predictions from algorithm 10; the difference is obvious with test series *K43*.

The generally good fits in Figure 4a are reflected in small mean squared deviations (MSD) of the predicted from measured data (Table 2). For DARK maximum targets (Table 2, cell A1), MSD values were smallest for

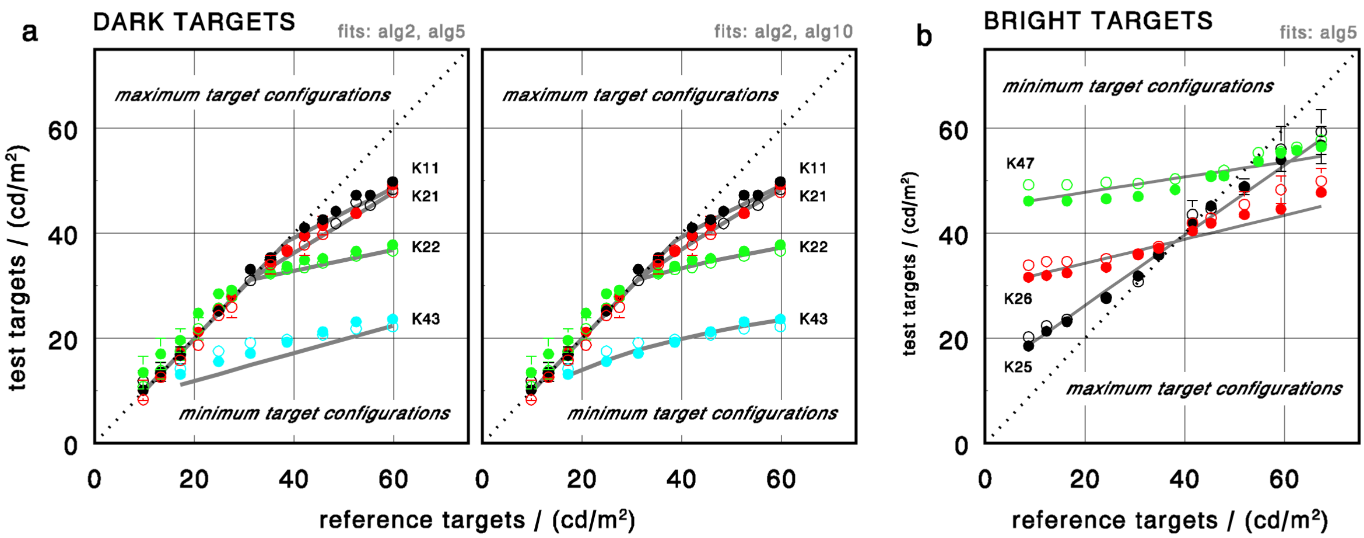


Figure 4. *Scatter plots of equal-salient target-target variations in Experiment 1.* Re-plot of the data in Fig.3 and an additional test series *(K21)* performed by the same three subjects. **a.** DARK targets; **b.** BRIGHT targets. Graphs show the luminance settings of equal-salient targets; matches from wide (open circles) and dense blob configurations (filled circles) are superimposed. Colors now distinguish data from different test series. Reference targets (abscissa) are matched by different test targets (ordinate) depending on the background luminance settings of each pattern (not shown); distractor settings were identical in these test series. Gray lines in each graph show the fits by the algorithms listed above each graph. For *DARK targets (a),* matches of maximum targets fall upon the identity line, *test target = reference target* (dotted), as predicted by algorithms 1-3 (cf. Table 1); matches of minimum targets in (lower right-hand part of the graphs) deviate from this prediction but are closely predicted by algorithm 5 (left-hand graph) and algorithm 10 (middle graph). For *BRIGHT targets (b)*, matches were best predicted by algorithm 5, but fits of predictions were considerably better for minimum targets (upper left-hand part of the graph) than for maximum targets (lower right-hand part).

algorithms 13, 2 (equivalent to 2a, 2b, and 3), and 1, and were generally larger for all other algorithms tested. For DARK minimum targets (Table 2, C1), MSD values were small for a number of algorithms including the two plotted in Figure 4a. (MSD values in Table 2 are based on *all* data from a given block of test series, not just those selected for illustration.)

For *BRIGHT targets* (Fig. 4b), best fits were obtained with the *salmin* algorithm 5 (Fig. 4b) and with algorithm 8b, both for maximum and minimum targets; the two fits were nearly identical for test series block *K* (Table 2, E1 and G1). Note however, that the best fits for BRIGHT maximum targets were generally not as good as the best fits for DARK maximum targets; MSD values are larger (cf. Table 2, E1 vs. A1) and the (best) gray-line predictions show larger deviations from the data (Fig. 4b), as was already seen in Figure 3b (test series *K26* and *K47*).

Conclusions from Experiment 1:

Target type	*Equal salience related to*
DARK maximum	constant-addition (alg1)
DARK minimum	salmin (alg5) and alg8b
BRIGHT maximum	salmin (alg5); poor fits
BRIGHT minimum	salmin (alg5)

Notable deviations of data from predictions with minimum targets (see Experiment 5)

Experiments 2 and 3: Different distractors on constant background *(Test series blocks L, LX, and O)*

In test series block *K,* distractors in reference and test patterns always had the same luminance. This restriction was avoided in test series blocks *L* and *O*, which were split to display only maximum or minimum targets.

Stimuli

In blocks *L* and *O* altogether 49 test series were run each with up to 9 test conditions; twenty-two of them were performed by all three subjects. In every test series, one constant reference pattern was compared with several test patterns at various distractor settings (cf. Fig. 5); background luminance was held constant and identical in the two patterns. Most test series were run with wide and dense blob configurations. However, salience matches with bright targets in sparse arrangements were generally more difficult than matches with dark targets and were hence not tested with all subjects in all conditions. To overcome the difficulties that seemed to occur with the adjustment of particularly bright targets (cf. Nothdurft, 2015), an additional test series block *LX* was created, in which test targets were dimmer than reference targets. Block *LX* included seven test series that were only tested in dense blob configurations.

Results and Discussion

Maximum targets: Test series blocks L and LX) *(Experiment 2)*

Figure 5 shows various examples of test series in blocks *L* and *LX*. Again, the results for DARK and BRIGHT targets differed systematically and, if tested, did not reveal strong differences between salience matches in dense and wide blob arrays.

DARK targets. Several straight-forward expectations (Nothdurft, 2015) are plotted into the data. One is based on the *constant-addition principle* ("add"; continuous gray lines), which assumes that targets have equal salience when they display the same luminance difference to distractors (algorithm 1 in Table 1). This principle provides one possible explanation for the data obtained with DARK maximum targets in Experiment 1, where distractor luminance was held constant (Fig. 3a). From Figure 5, however, it is obvious that this principle cannot explain the experimental data of all test series when distractor luminance is varied. In particular, targets close to the monitor limit were adjusted brighter than predicted (*L12, L22)*. The second expectation plotted in Figure 5 is based on the *constant-ratio principle* ("ratio"; dashed gray lines); it assumes that the ratio of target and distractor luminance must be constant to let targets appear equally salient (algorithm 2 in Table 1; Weber contrast). This principle, too, can explain the matches of DARK maximum targets in Experiment 1. In Figure 5a, the constant-ratio principle predicts the experimental data in some curve sections (*L22, L51)* but fails in others. The third prediction is from the *salmin* algorithm 5 (black thin lines); it clearly misses the data for DARK targets in Figure 5.

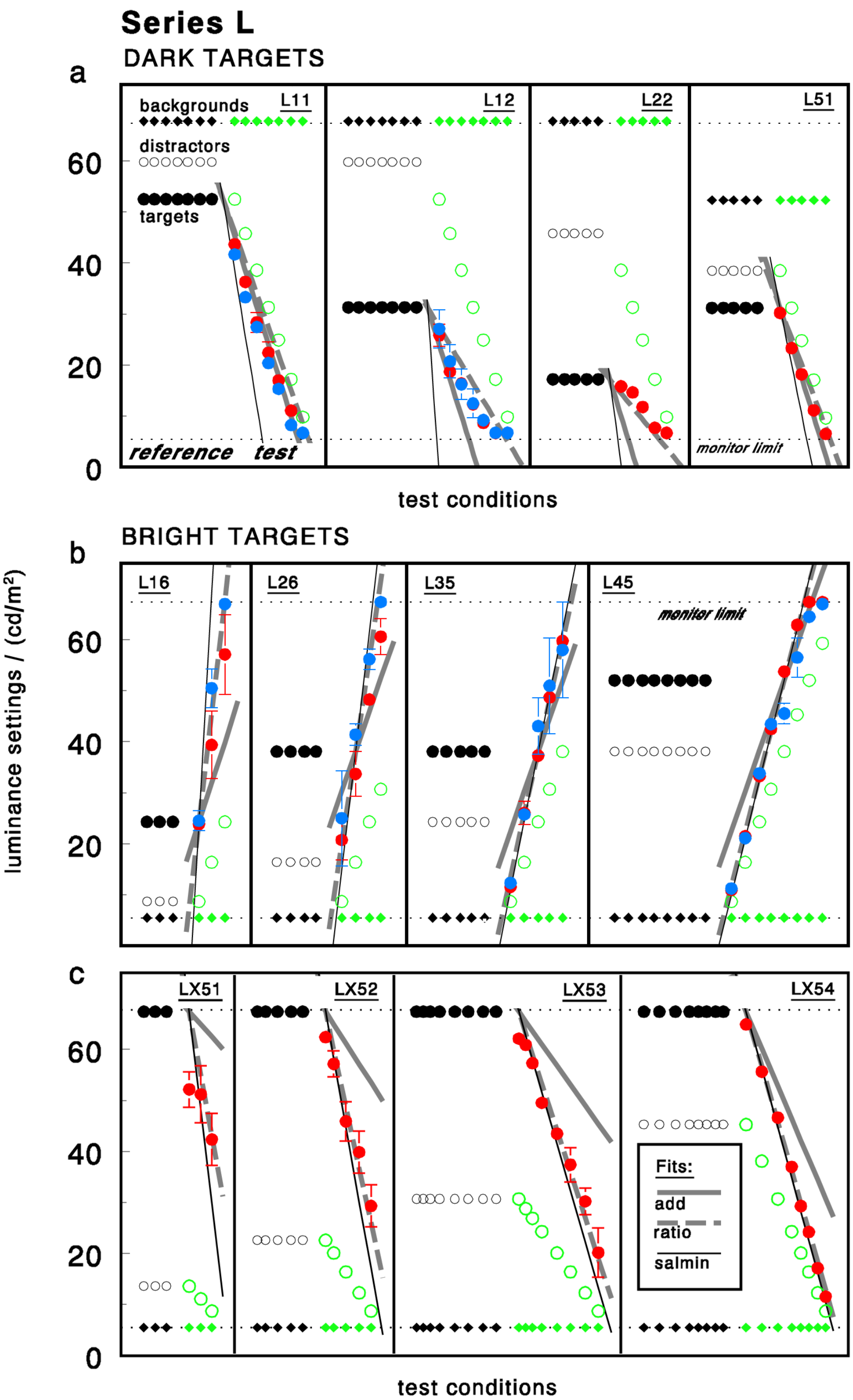


Figure 5. *Luminance settings of salience-matched target conditions in Experiment 2 (test series block L; all targets in maximum configuration).* **a.** DARK targets; **b, c.** BRIGHT targets; symbols and data presentation as in Fig.3. In these test series, reference patterns (black symbols) were held constant, and test distractor and target settings were varied; backgrounds were identical in the two patterns. Straight gray and black lines show various predictions of the experimental data. Of the predictions shown, the constant-ratio prediction fitted all data best. Thin dotted lines indicate the upper and lower luminance settings of the monitor. Test series *L (b)* and *LX (c)* differed in the relative brightness of the test targets to be adjusted. In *(b)* they could be the brightest items in the scene which has sometimes led to difficulties in adjustments (not in the series shown here). In *(c)* that was never the case.

BRIGHT targets. The corresponding fits for BRIGHT targets are shown in Figure 5b and c. Since matches sometimes varied considerably (cf. large s.e.m.) when test distractors were brighter than reference distractors and thus the targets to be adjusted were the brightest items in the display (cf. Anderson, Singh, & Meng, 2006; Nothdurft, 2015), an additional test block series *LX* was created in which test distractors were dimmer than reference distractors and test targets did not have to exceed all other items' luminance.

In both test series blocks, the constant-ratio principle (dashed gray lines) fitted the data much better than the constant-addition principle (continuous gray lines); the difference is particularly obvious with data from test series block *LX* (Fig. 5c). The third class of curves, predictions from the *salmin* algorithm 5 (thin black lines), often fell close to the data but generally not as close as predictions from the constant-ratio principle.

Underlying algorithms

DARK targets. The limited validity of certain algorithms to predict the data of Experiment 2 is also seen in the scatter plots of Figure 6a. Since reference patterns were constant in each test series and backgrounds were identical in reference and test patterns, the individual matches (circles) together with the reference condition of a series (squares) should show a systematic variation of target luminance with various distractors. The two left-hand graphs replicate two predictions already shown in Figure 5a; constant addition (gray lines in the left-hand graph of Fig. 6a) and constant ratio (gray lines in the middle graph). None of these predictions fits the entire data set. This excludes Weber and Michelson Contrast (algorithms 2 and 3) as salience algorithms for DARK maximum targets when distractors are not identical. To search for better fits, several algorithms from Table 1 were tested (section C). The best fit was obtained with algorithm 8 (Fig. 6a, right-hand graph), the normalized Michelson Contrast (Singh & Anderson, 2006).

The different quality of the fits is reflected in the MSD values of the three predictions. The smallest MSD is obtained for algorithm 8, whereas algorithms 13 and 2 (the

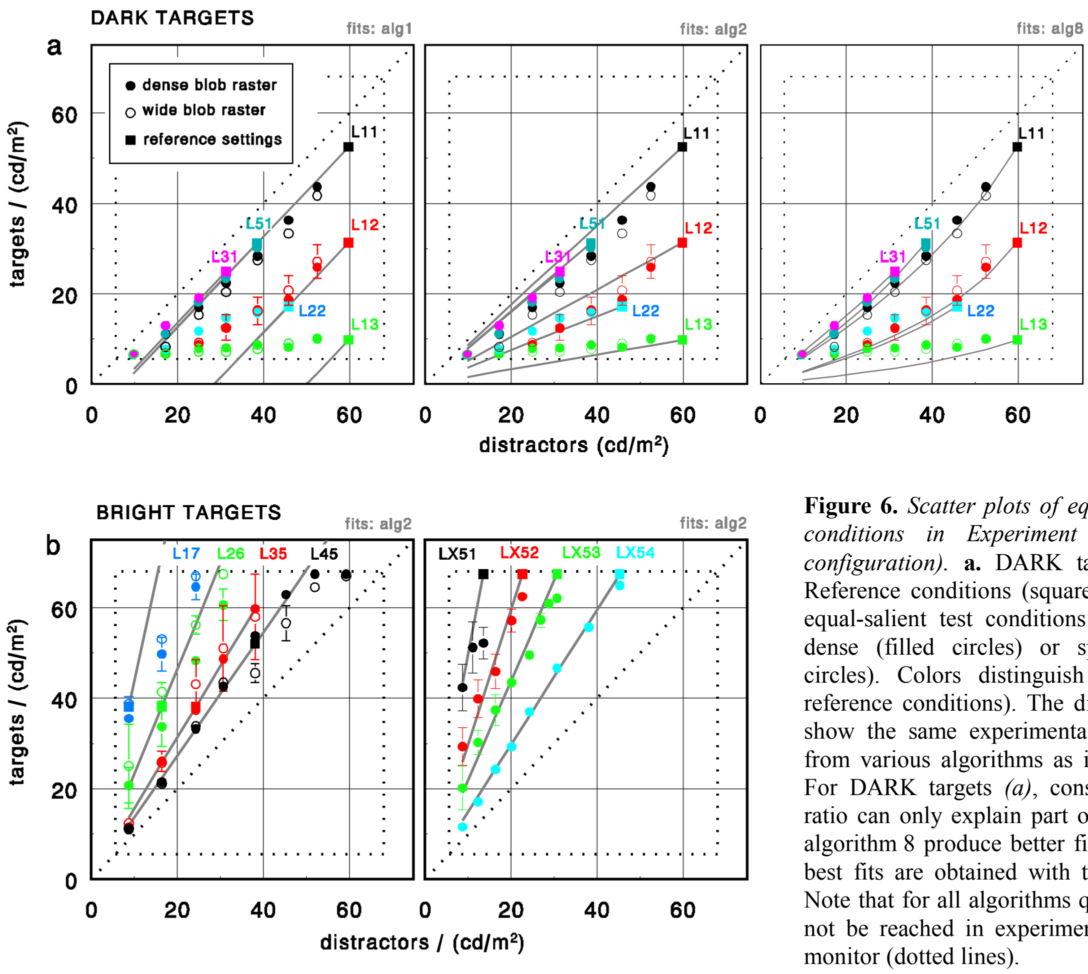


Figure 6. *Scatter plots of equal-salient target-distractor conditions in Experiment 2 (targets in maximum configuration).* **a.** DARK targets; **b.** BRIGHT targets. Reference conditions (squares) are shown together with equal-salient test conditions obtained from matches in dense (filled circles) or sparse blob patterns (open circles). Colors distinguish test series (and different reference conditions). The different graphs in *(a)* or *(b)* show the same experimental data fitted by predictions from various algorithms as indicated above each graph. For DARK targets *(a)*, constant-addition and constant-ratio can only explain part of the data; predictions from algorithm 8 produce better fits. For BRIGHT targets *(b)*, best fits are obtained with the constant-ratio algorithm. Note that for all algorithms quite a few predictions could not be reached in experiment due to limitations of the monitor (dotted lines).

best ones for test series block *K*) generated larger deviations (Table 2, A2). Algorithm 1 is not listed among the five best ones because too many predictions fell outside the monitor range (see section C).

BRIGHT targets. While Weber and Michelson Contrast (algorithms 2 and 3) could not predict the equal-salience settings of DARK targets, they better predicted the matches of BRIGHT targets in Figure 6b. Note however that not all test series were adequately fitted by these algorithms. For some series, in particular series *L17*, predictions fell far off the data points. This was much better with test series block *LX* (right-hand graph Fig. 6b) of which almost all data points lie close to the predicted curves. The main difference between these two test series was the brightness ranking of test targets. In some patterns of series *L*, the test target was the brightest item and had to be adjusted to values above all other luminance settings in the stimulus. In test series *LX*, the brightest item was always the reference target, which remained constant during adjustments. Obviously, adjustments within a given luminance range could be better performed and led to more reliable results than adjustments outside and above that range (Anderson, Singh, & Meng, 2006; Nothdurft, 2015), for example, in test series *L17* and *LX51* (Fig. 6b).

In spite of these differences it may be surprising that the MSD values from test series *L* and *LX* are so similar (Table 2, E2). But this is misleading. The few data points in the left-hand graph (test series *L17)* that lie far off the predictions were not included in the MSD computation, since constant-ratio predictions fell outside the available monitor range (for details of which data points were included and which not, see section C). The remaining data of test series block *L* are closely fitted by predictions from algorithm 2, as are the data from test series block *LX*. It is interesting to note that also the *salmin* algorithm 5 made good predictions of series *L* (MSD value 19.1) but is not listed in Table 2, E2 because too many predictions fell outside the monitor range.

Conclusions from Experiment 2:

Target type	*Equal salience best related to*
DARK maximum	normalized Michelson (alg8)
BRIGHT maximum	salmin (alg5)

Some algorithms had to be ignored because too many predictions fell outside the available monitor range. – Problems with very bright targets.

Minimum targets: Test series block O (Experiment 3)

Examples of salience matches with test series block *O* are shown in Figure 7. All patterns in these series presented targets in the minimum configuration. Like in series *L*, reference patterns were held constant within a series, and in the test patterns only distractor and, of course, target luminance settings were varied. Background settings were identical in reference and test patterns.

For *DARK targets* (Fig. 7a), neither the constant-addition (continuous gray lines) nor the constant-ratio principle (dashed gray lines) can predict the entire data set, although predictions partly follow the data. This is particularly obvious with matches in wide blob configurations (blue symbols) of test series *O12*. The data follow the *salmin* prediction when the test target was brighter than the reference target, but follow more closely the *constant-addition* scheme when the test target was darker than the reference target. Similar differences can be seen in the other tests series of block *O*. For *BRIGHT targets* (Fig. 7b), all data are closely met by the *salmin* prediction (algorithm 5; thin black lines), but differences between algorithms are not very pronounced in some test series (e.g., test series *O48* and *O69).*

In computational predictions of the *DARK target* data, algorithm 10 produced the smallest MSD values (Table 2, C2) and the visually best fit of the data in the scatter plots of Figure 8a (left-hand graph). The suspicion that the data might, in fact, be fitted by two algorithms, the *salmin* prediction (when the test target is brighter than the reference target) and the constant-addition principle (when the test target is darker than the reference target), is shown in the right-hand graph. But although these bipartite fits are sometimes close to the data, there are notable deviations (e.g., test series *O11*) which do not occur in the fits from algorithm 10.

For *BRIGHT targets*, several algorithms produced rather good predictions of the data (Table 2, G2). Nearly perfect fits for wide and dense blob configurations were obtained with the *salmin* algorithm 5 plotted in Figure 8b.

Conclusions from Experiment 3:

Target type	*Equal salience best related to*
DARK minimum	algorithm 10 or bipartite fits (but see Experiment 6)
BRIGHT minimum	salmin (alg5)

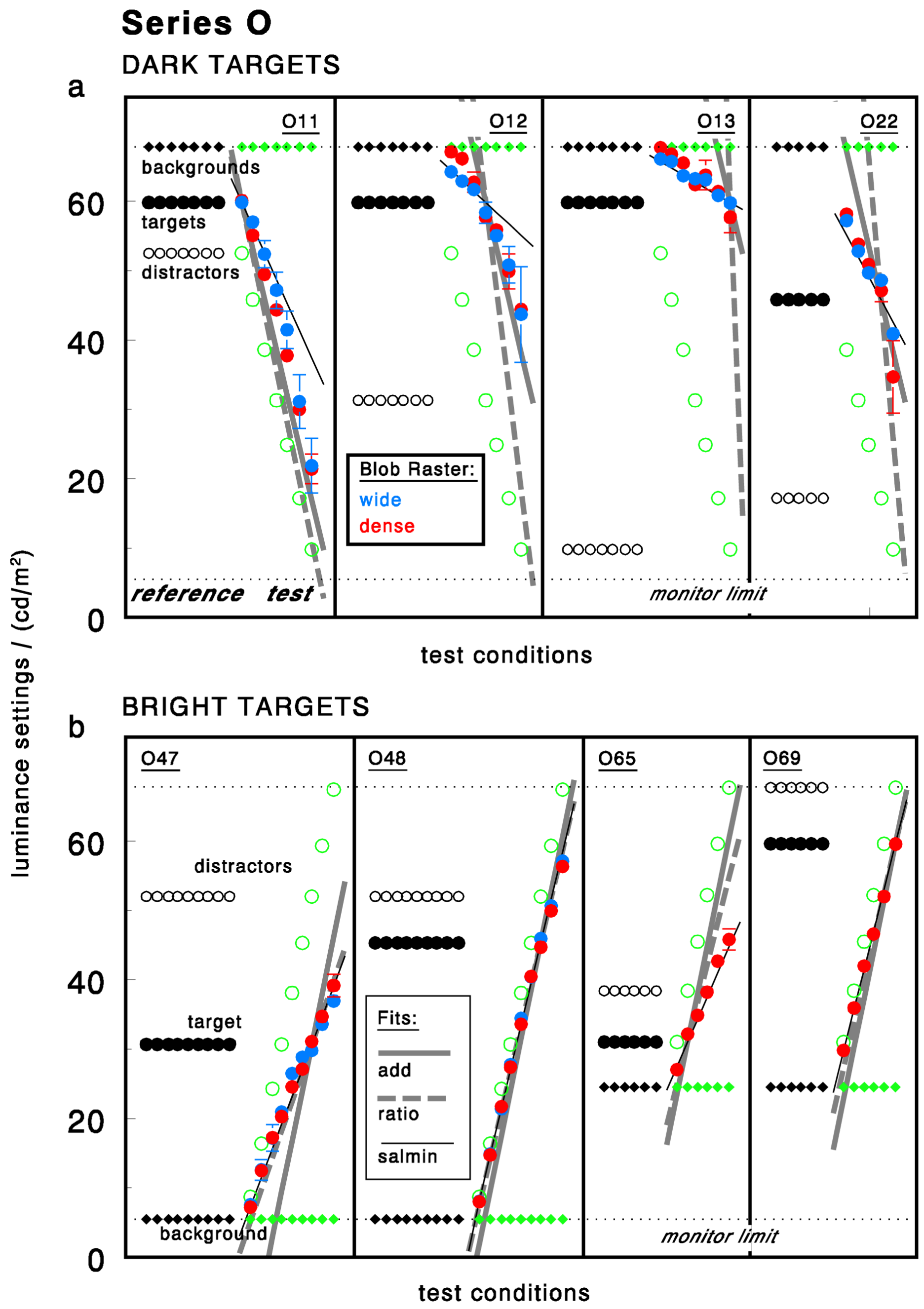


Figure 7. *Luminance settings of salience-matched target conditions in Experiment 3 (test series block O; targets in minimum configuration).* **a.** Examples of test series with DARK targets; **b.** examples with BRIGHT targets; symbols and data presentation as in Fig. 5. With DARK targets *(a)*, some curves appear to be split; the different branches follow different algorithms ("bipartite fits"). With BRIGHT targets, best fits were obtained with the *salmin* algorithm.

Experiment 4:
The effect of background variations
(Test series block F)

We have seen in Experiment 1 that background luminance could strongly affect the salience of certain targets but had almost no effect on the salience of DARK maximum targets (Fig. 3). To consolidate this observation, an additional block of test series was designed, in which reference and test patterns were identical and only the background in the test patterns was varied between test conditions. If background luminance would affect

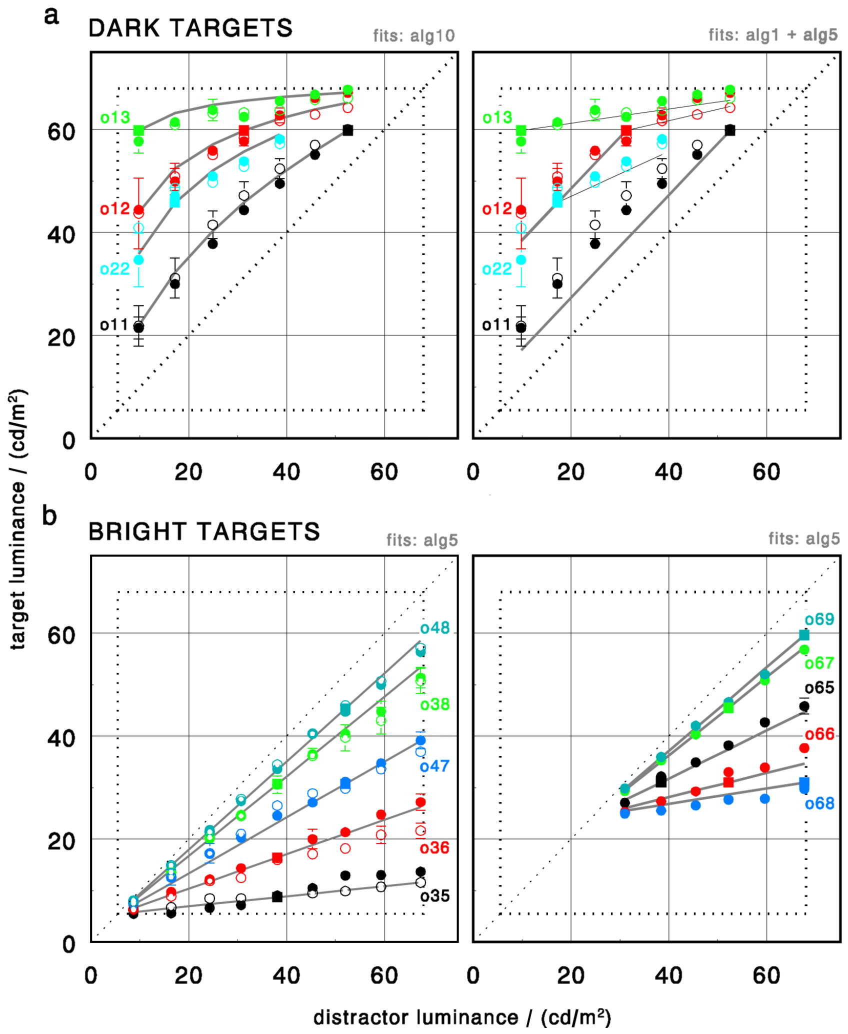


Figure 8. *Scatter plots of equal-salient target-distractor conditions in Experiment 3 (targets in minimum configuration).* Symbols and presentation as in Fig. 6. **a.** DARK targets; the same data are fitted by algorithm 10 (left-hand graph) and, in "bipartite fits", by algorithms 1 and 5 (right-hand graph). The latter fits do not follow test series *O11*. **b.** BRIGHT targets. All data are closely predicted by algorithm 5.

salience, test target settings should vary. If background luminance is irrelevant for salience estimates, target adjustments should be identical between test conditions.

Stimuli

All major target conditions tested so far (DARK and BRIGHT targets in maximum and minimum configuration) were studied in altogether ten test series (Fig. 9). In each series, the luminance settings of the reference pattern and of the test distractors remained constant; only the test pattern background was varied. In every presentation, the test target had to be adjusted to match the salience of the reference target, as in the previous experiments. Three subjects performed the tests with dense blob configurations, one subject in addition with wide blob configurations.

Results and Discussion

Since reference patterns and test patterns were identical except for the different background settings in the test patterns, the variations in salience-matched test target luminance should directly reflect the influence of background luminance on target salience. For *DARK targets in maximum configuration* this influence was indeed very small (Fig. 9a, red and dark-blue symbols). Only when background luminance closely approached that of distractors so that their visibility was strongly diminished, subjects tended to diminish also the *target*

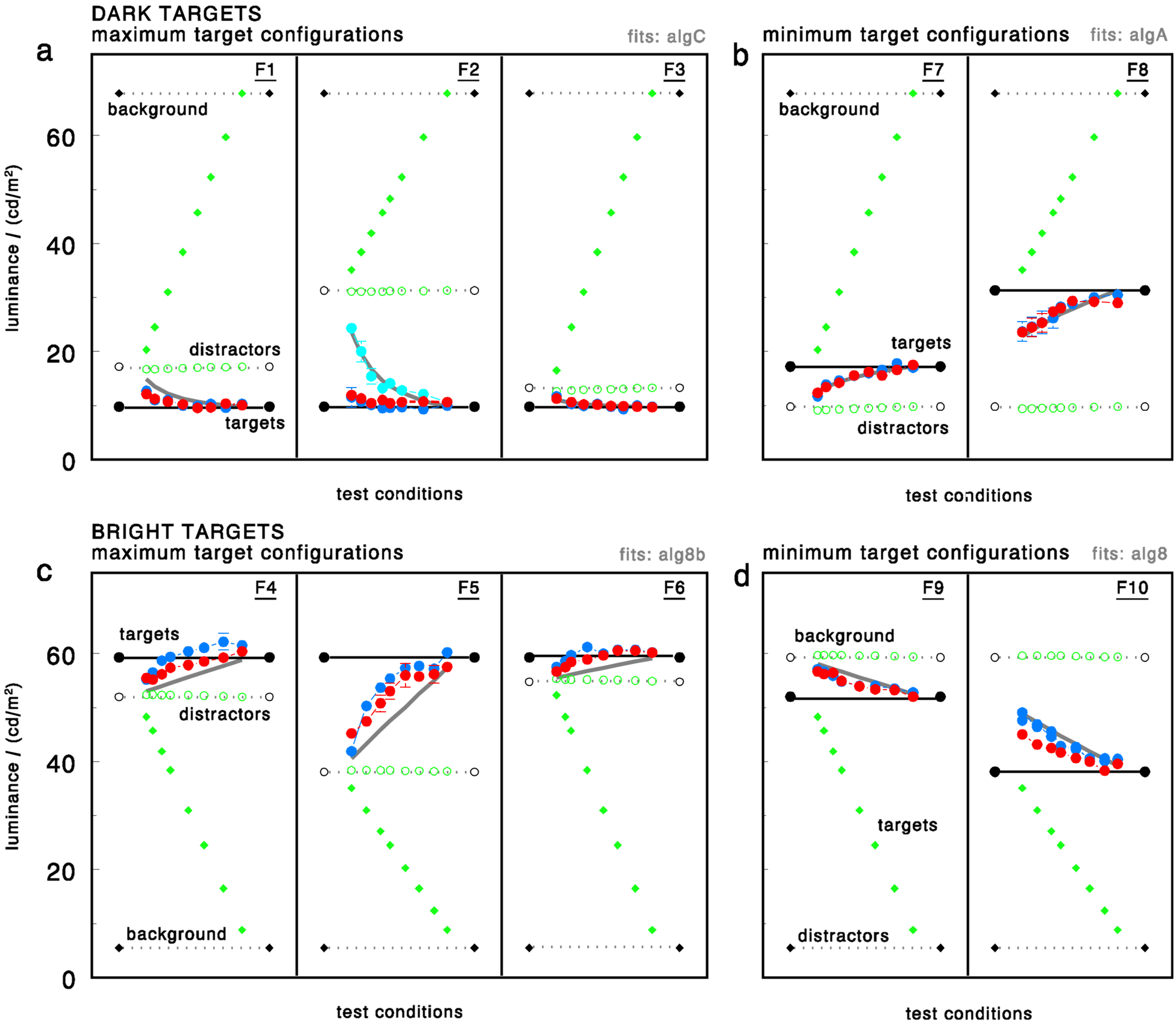


Figure 9. *The influence of background luminance on equal-salience matches (Experiment 4, test series block F).* **a.-d.** Different test series for DARK *(a, b)* and BRIGHT *(c, d)* targets in maximum *(a, c)* and minimum configuration *(b, d)*; symbols and data presentation as before. In each test series, a fixed reference pattern was matched by test patterns with same distractors but different backgrounds. If background settings were irrelevant, matched test targets should all fall upon the line of the (constant) reference target. Deviations from this line directly indicate background effects. Gray superimposed lines show predictions from the algorithms (cf. Table 1) listed above the plots of each target group. Best predictions for DARK maximum targets were from algorithms 1, 2, and 13, which all predict straight lines superimposed on the "targets" lines (not shown). In wide blob arrangements of test series *F2*, different matches were obtained when target contrast (light blue circles), not global salience (dark blue circles) was matched. This performance was fitted by algorithm C.

contrast to background and thus effectively the salience-matched target-to-distractor contrast. This confirms and extents the finding of Experiment 1 that the salience of DARK targets in maximum configuration mainly depends on the target-to-distractor contrast, unless the distractors themselves are barely discriminable from background. An exception was seen with medium distractor levels in *wide* blob configurations (test series *F2,* blue circles). Here different matches were found depending on which aspect of the stimulus was paid attention to. When test targets were matched for similar *contrast to distractors* (light blue data points), they were adjusted brighter than when matched for global similarity in salience (dark blue data points). No such difference was found in *dense* blob configurations. The deviations were unexpected and further explored in a spontaneous control test with short stimulus presentations (500 ms). Only the settings with dark blue symbols in Figure 9a let targets appear about equal-salient to the reference targets (black). When targets with the light-blue luminance settings were compared in such a test, the reference targets were generally considered to be more salient.

In *all other stimulus conditions* (Fig. 9b-d) variations of background luminance had a much stronger effect on target salience, and in many tests target luminance had to be changed considerably to make the targets on different backgrounds equal-salient. Performance was generally quite similar in wide and dense blob patterns.

Underlying algorithms

Except for BRIGHT targets, in particular those in maximum configuration, all experimental data are closely predicted by the algorithms listed above each graph (gray lines in Fig. 9). *DARK maximum targets* are, as in Experiment 1, perfectly fit by algorithms 1, 2 (=3), and 13 (MSD values $\leq 2.0^2$ in Table 2, A3), which predict no or only small variations with background. These predictions fall upon the horizontal line “targets” and are not plotted in Fig. 9a. Contrast-based matches (light-blue data points in *F2)* are closely predicted by algorithms A and C (MSD values 1.6 and 3.1, respectively; the values for these special matches are not listed in Table 2, A3); the predictions from algorithm C are shown in Figure 9a. *DARK targets in minimum configuration* (Fig. 9b) are best predicted by algorithm A (Table 2, C3). The fits to BRIGHT targets are generally worse. Among the best fits (smallest MSD values) for *BRIGHT maximum targets* (Table 2, E3) are predictions from algorithm 8b (plotted in Figure 9c) and from the *salmin* algorithm 5; however, the deviations from the experimental data are quite obvious (a better fit will be presented in section C). For *BRIGHT minimum targets* (Table 2, G3), the best fit was obtained with algorithm 8 (plotted in Fig. 9d); it was quite good for matches in wide but not dense configurations.

2 Note that all MSD values in this paper have the physical unit (cd/m^2)^2, which is left out for the brevity of presentation. See General Methods.

Conclusions from Experiment 4:

Does salience vary with background luminance?

DARK maximum	(almost) no
DARK minimum	yes
BRIGHT maximum	yes
BRIGHT minimum	yes

Experiment 5: Variations in minimum target matches *(Test series block R)*

While salience matches of targets in minimum configuration by and large followed the *salmin* predictions (cf. series *K11, K22,* and *K25* in Fig. 3), there were sometimes strong deviations from the straight prediction curves when test targets had to be adjusted in a rather small distractor-background window (cf. series *K43, K26,* and *K47* in Fig. 3). These deviations appeared to be particular strong in patterns with wide blob spacing. Experiment 5 was designed to explore these variations in more detail.

Stimuli

The experiment was added at a late stage of the project, after two subjects of the main study had left. They were replaced by three new subjects (two male, one female) who had not been involved in the other experiments of the study. The author has served as an additional observer in all experiments.

Experiment 5 included four test series with minimum targets, two with DARK and two with BRIGHT targets. All matches were performed on patterns with *wide* blob arrangements. Each test series included 19 test conditions that were presented twice in a pseudo-random sequence. Test series were designed to study the transform of minimum target variations in a large background-to-distractor span, into equal-salient minimum target variations in much smaller background-to-distractor luminance spans (Fig. 10). Within a series, background and distractor luminance settings were held constant (different for reference and test patterns) and only target luminance was varied. Across the series, two different background and distractor settings were tested. In every single trial, the test target had to be adjusted to match the salience of the reference target, as in all previous experiments.

Results and Discussion

The results are best visualized in scatter plots of equal-salient reference and test targets (Fig. 10); the constant luminance settings of background and distractors are indicated by straight lines. In test series *R1* and *R4,* the backgrounds in the two patterns (continuous lines) were identical and distractors (dashed lines) differed; in test series *R2* and *R3,* distractors were identical and backgrounds differed. In all test series, the luminance span of backgrounds to distractors was much larger in the reference than in the test patterns. It was the aim of Experiment 5 to study if and how these different spans affected the matches of equal-salient targets.

If salience were strictly scaled to the luminance span of background and distractors, as originally assumed, all data points should fall upon the lines predicted by the *salmin* algorithm 5 (thin black lines). But this was not the case. Instead, equal-salience matches let targets often shift away from these predictions. The deviations mainly occurred when target luminance was either close to distractor luminance (near the corners made up by dashed lines) or close to the backgrounds (near the corners made up by continuous lines), and are quite different in the four test series. In all corners, however, luminance variations of the test target (embedded in the smaller luminance span) tended to exceed the luminance variations expected from the *salmin* algorithm. To make the test targets equal-salient to the reference targets, subjects had slightly increased the luminance difference to either background or distractors. Farer away, in the middle of most curves, matches cross the *salmin* predictions.

It is interesting to compare these deviations with other predictions. For targets very similar to the background (corners of continuous lines), all matches follow the *constant-ratio* principle (gray dashed lines). This was not the case for targets similar to distractors (corners of dashed lines). Constant-ratio predictions are not shown at all these

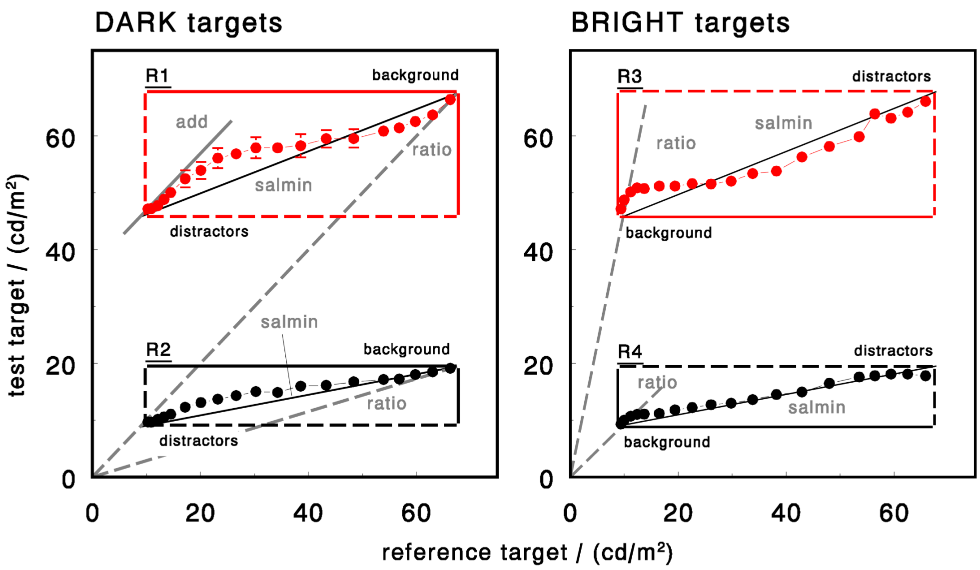


Figure 10. *Deviations from salmin predictions for minimum targets (Experiment 5, test series block R).* All tests were performed on wide blob arrangements. Graphs plot luminance variations of equal-salient targets in test series *R1-R4*. In each series, background and distractor settings were constant (indicated by straight lines; backgrounds continuous, distractors dashed). For small distractor-to-background differences, which were tested here, equal-salient target matches were not linearly related to the luminance span of background and distractors (as predicted by the *salmin* algorithm; thin black lines) but showed strong local deviations towards other computational rules (gray lines). Particularly important was the constant-ratio principle (dashed gray lines) for small target-to-background differences (data points near the corners formed by background luminance settings) but not for target-to-distractor differences (data points near the corners of distractor settings). Of the other simple predictions tested here only the constant addition of target-to-distractor differences was partly important for DARK targets. For the construction of further predictions note that constant-ratio curves connect the origin with the points of interest (i.e., background corners for target-to-background differences; distractor corners for target-to-distractor differences), whereas constant-addition curves run through these points in parallel to the diagonal identity lines, *test target = reference* (dashed gray line in the left-hand graph).

corners in Figure 10 but can be easily transferred from the neighboring graphs; they form straight lines between the corners and the origin. It is important to notice that the curves through equal-background corners (labeled "ratio") do *not* represent the Weber contrast of targets and distractors as given in algorithm 2 of Table 1, but the Weber contrast of targets to *background*. The fits thus indicate that equal-salient targets close to background vary in proportion to the background luminance (Nothdurft, 2015). The missing fits at the opposite equal-distractor corners (where curves indeed represent constant target-to-*distractor* ratios as given by algorithm 2 in Table 1) simultaneously indicate that algorithm 2, the Weber contrast of targets and distractors, cannot explain the salience of targets in minimum configuration. Predictions from *constant addition* (continuous gray lines), on the other hand, did only fit the matching data of DARK targets similar to distractors (series *R1)*. For clarity, these predictions are not drawn into all graphs but can again be easily constructed; constant-addition predictions are represented by straight lines parallel to the main diagonals of the two graphs *(test target = reference target)* through any of the given corners. For corners that lie upon the diagonals (backgrounds in *R1* and *R4*; distractors in *R2* and *R3)*, constant-addition and constant-ratio predictions are identical.

While the local deviations from *salmin* and the alternative fits to constant-ratio or constant-addition predictions are quite obvious in Figure 10, there were individual variations between subjects. They mainly occurred in the middle of the windows (where targets were similarly different from both background and distractors) and partly at the dashed lines' corners (targets similar to distractors). Here, some subjects followed better than others the *salmin* rule, in some test series. Such variations were, however, small when targets were close to background luminance, i.e. near the corners of continuous lines. Only one of the four subjects (and in only one test series) did *not* strictly follow the constant-ratio principle at this end.

Altogether, we must thus restrict the previous conclusion that targets in minimum configuration follow the *salmin* algorithm. This conclusion is only valid in a first, global overview. When minimum targets are close to background and presented in a small luminance window of background and distractors, they more closely follow constant-ratio settings of targets to backgrounds.

Conclusions from Experiment 5:

In matches of *minimum targets*, strong deviations from the *salmin* prediction will occur in wide patterns with small distractor-to-background luminance variations. Differences between targets and distractors and, in particular, between targets and background may then be enhanced.

Experiment 6: Bipartite fits of DARK minimum target matches?

We have seen in Figure 7a that matches of a constant DARK minimum reference target appeared to follow the *salmin* prediction when test distractors were brighter than reference distractors, but seemed to follow the constant-addition rule when test distractors were darker than reference distractors. This first impression was only partly confirmed in Figure 8. Some test series were indeed closely predicted by bipartite fits but others were not and better predicted by algorithm 10. To further explore the phenomenon, new test series were added to block *O,* in which the presumed switching from one computational rule to the other was studied, in some series with much better resolution.

Stimuli

Stimuli were similar to those used for DARK targets in test series block *O* (Experiment 3); that is, all targets were presented in the minimum target configuration, and within a test series the reference pattern was held constant while test distractors were varied. The backgrounds of test and reference patterns were identical.

The additional tests were performed in two steps. First, a large number of additional tests series was explored, some of which are presented in Figure 11. These series were studied in dense and wide blob configurations. In a second step, three series were chosen in which the presumed switching between different rules should have been visible, and these series were then expanded to display the presumed performance switching in much better resolution. Each of these test series contained 13-15 conditions that were repeatedly tested in random order. In addition to the wide blob raster stimulus patterns were also presented in a new *"packed"* presentation in which targets immediately touched the distractors, and these their neighboring distractors, so that the background could not be seen. Only two luminance levels were present in these latter patterns, *tg* and *dis*. Furthermore, in one of these series (O68) matches were performed in two different modes, in the standard mode with alternating foveal inspection of the two targets and in a “global inspection” mode in which the targets were matched from a gaze position in the middle between the two patterns (cf. Exp. 4 in Nothdurft, 2015). The (time-consuming) tests were run by subject HCN.

Results and Discussion

Since the experiment was a follow-up of Experiment 3, certain findings must be discussed together (see below). Here, I will only briefly present a selection of the main observations.

Figure 11 shows the matches of six of the additional test series. In addition to the mean performance (Fig. 11a) also the individual matches from repeated trials are plotted at each test condition (Fig. 11b). Note that in only one test

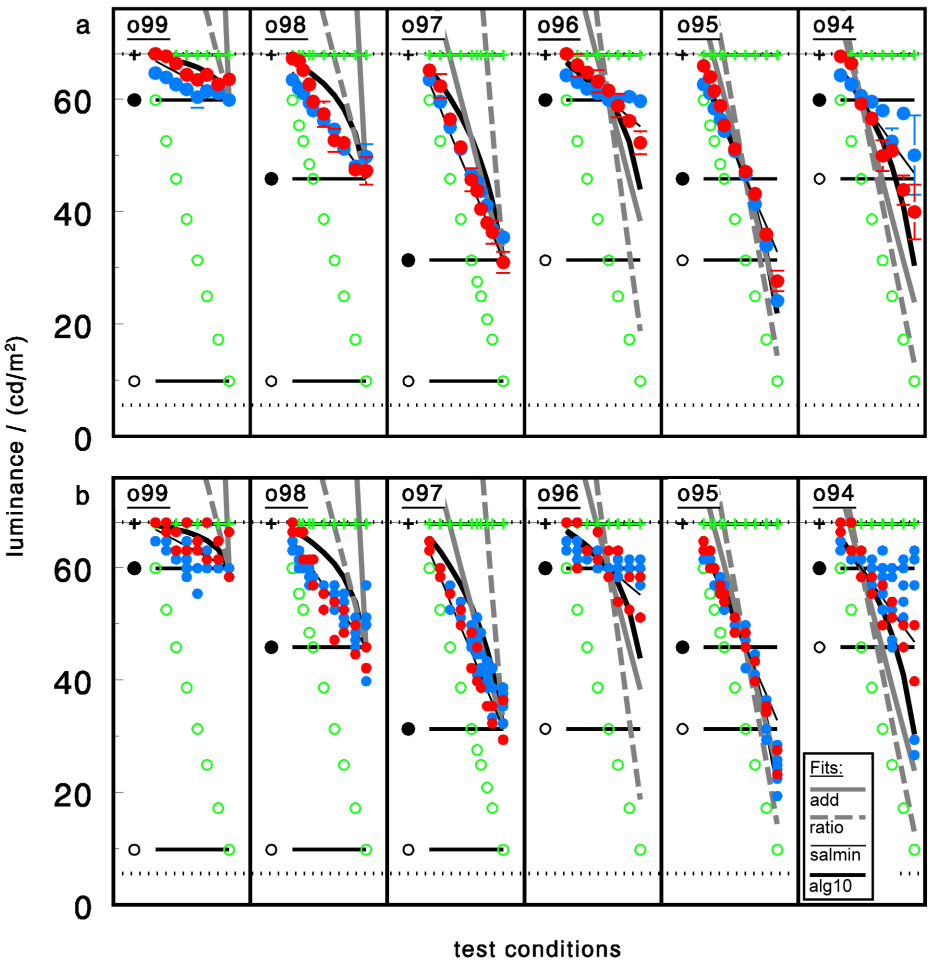


Figure 11. *Additional test series in block O (Experiment 6).* Graphs plot equal-salience matches; **a.** means; **b.** individual trials. All patterns had the same backgrounds (crosses). Within a test series, reference patterns were constant (black symbols and horizontal lines) and test distractors (open green circles) were systematically varied. Blue data points represent test target matches obtained in the wide blob raster, red data points matches in the dense raster. Gray and black line curves show predictions from three standard algorithms (as indicated) and from algorithm 10 (thick black lines). In certain conditions, matches were less certain than in others and single matches *(b)* strayed widely *(O94)*. In only one test series *(O95)* did matches follow the predictions from algorithm 10.

series *(O95)* data followed the predictions from algorithm 10 (thick black lines); in all other series the data clearly deviated from these predictions. Instead, matches strictly followed the *salmin* predictions when test distractors were brighter than reference distractors (green open circles above the black open circle level), as in the entire curves of test series *O99*, *O98*, and *O97,* but were less predictable when test distractors were darker than reference distractors (right-hand parts of test series *O96*, *O95*, and *O94)*. This is particularly obvious in test series *O94*, where the single trial matches (Fig. 11b) in these conditions were widely scattered.

To study this phenomenon in more detail, in particular at test conditions where one rule may change into another one, three final test series with better resolution (i.e. with finer gradation of distractor variation) were run. Figure 12 again shows mean performance (Fig. 12a) and the individual matches from repeated trials (Fig. 12c). Series *O12a* resembles the luminance setting of the original series *O12,* in finer resolution; the other two series are new. Matches of targets in wide blob arrangements are shown in blue, the additional matches of targets in packed configurations in purple.

It is interesting to look at the distribution of individual matches (Fig. 12c) in different test conditions of each graph. While the individual matches in wide blob arrangements (dark-blue data points) fall relatively close together when test distractors were brighter than the reference distractors (in the left-hand parts of each graph), they tend to scatter more widely when test distractors were darker than the reference distractors (right-hand parts of each graph). This larger scatter (also seen in test series *O94* in Fig. 11) might have been caused by several effects. Some targets were still matched according to the *salmin* algorithm (thin black lines); other matches were shifted more closely to the constant-addition rule (gray continuous lines). And in some matches, test targets were apparently even matched to display the same luminance as the according reference targets (data points near the horizontal line). These latter matches were the preferential performance in the global matching mode (targets were simultaneously matched from a central fixation point between the patterns) when test distractors were darker than the reference distractors (Fig. 12c and d). This uncertainty of which match to follow apparently produced larger variations for matches with test distractors outside the reference background-to-distractor window than inside that window.

Also quite interesting are the results from matches with blobs in *"packed"* configurations. Because the background is not visible in these patterns (target and distractors are touching each other), the DARK minimum targets are, in fact, BRIGHT single targets on a distractor luminance background. Equal-salience matches of such targets should follow either the constant-ratio or the constant-addition rule (Nothdurft, 2015), indicated here by continuous and dashed gray lines, respectively. The true matches lay exactly between these predictions so as if neither rule was particularly stringent and the subject could not decide which one to follow.

Note that the data of Experiment 6 neither support the strict bipartite model nor the fit from algorithm 10 (thick black lines). When the test distractor was darker than the reference distractor, curves could follow different rules; they could, at least for some test conditions, continue to follow the *salmin* prediction *(O12a, O19, O68)* or widen the range of target adjustments to darker or brighter luminance settings *(O94, O68),* and might only finally switch to the constant-addition rule when distractors are sufficiently different. At which distractor difference the decision changes, may depend on the mode of how subjects performed the matches and on the bias of the entire sample of test conditions in a given test series. When biased to compare the brightness of targets or to follow the *salmin* algorithm, quite a few matches with darker test distractors might still be adjusted too high. When instead biased to look for similar target-distractor differences, matches might more often follow the constant-addition rule. This bias should naturally be also affected by the relative frequency of test conditions in a particular run. When a test series contains many test conditions that clearly follow the *salmin* rule, a few uncertain matches might also be adjusted this way. And when a test series contains many conditions that force adjustments according to the constant-addition rule, the uncertain matches might be biased in this direction.

Experiment 6 has provided new observations and interpretations which I should like to summarize before going into a general discussion of the findings so far. First, the means of equal-salience matches do not necessarily represent the mean of a single algorithm but may represent the averaged performance of two (or more) different algorithms. Second, equal-salient target matches may follow different rules depending on the relative luminance settings of reference and test distractors. Which rule an observer would follow at transitions between these rules,

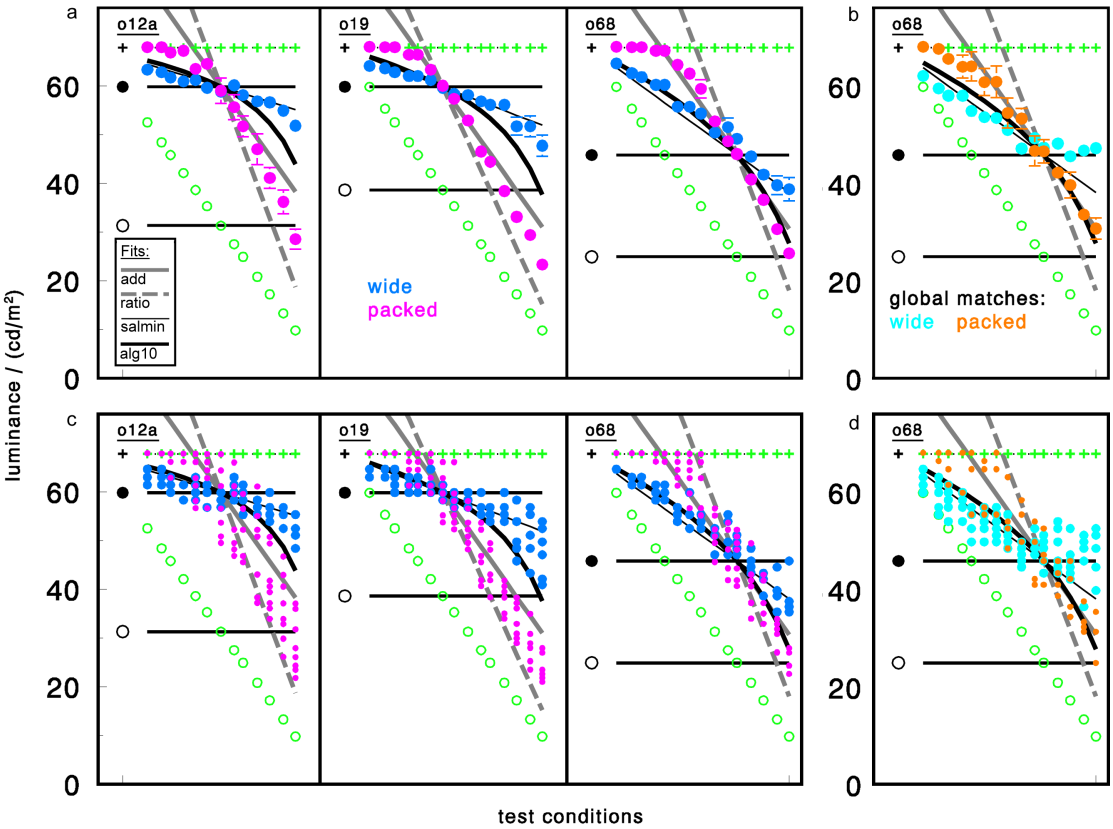


Figure 12. *Test for bipartite fits in test series block O (Experiment 6).* Graphs plot equal-salience matches in three test series with high resolution; **a, b.** means; **c, d.** individual matches; presentation as before (backgrounds as crosses). Blue data points represent matches from wide blob arrangements; purple and orange data points are from new matches with *"packed"* blob arrangements in which targets and distractors touch each other so that the background is not seen. Matches were obtained in alternating foveal target inspections (a, c) or in "global inspections" from a fixation point between the patterns. Gray and black line curves show the indicated predictions. None of these test series provided evidence for bipartite fits of the data, and in none are the data fitted by algorithm 10. Note that matches in the packed blob arrangements fell strictly between two standard predictions (purple and orange data points).

may be uncertain but may likely also depend on the relative frequency of test conditions in the sample that clearly follow one rule. Third, complex algorithms (like algorithm 10, thick black lines in Fig. 11 and 12) may successfully predict the combined performance of different rules in certain tests but may fail if transitions are shifted.

Conclusions from Experiment 6:

For certain targets (here DARK minimum), salience matches depend on the ranking of test and reference targets. In transition regions, different rules may coexist and leave the observer uncertain about which rule to apply. Preferences may also depend on the test sample.

Discussion of Section A

Hints for reading: The Discussion first summarizes the results so far, stresses certain inconsistencies therein, and puts the observations into the context of luminance variations in nature. It may be worth reading. Headed sections further down address various details of the observations, which you may skip if not interested. Important is the distinction of item salience and discrimination salience. Look at the demos in Figures 13 and 14. There is a summary at the end of this Discussion section.

Dedicated to uncover the computation of luminance-defined salience of targets among distractors, Experiments 1-6 revealed several important findings. First, the salience of various targets is represented by different mechanisms; that of DARK targets was computed differently from that of BRIGHT targets and that of maximum targets sometimes differently from that of minimum targets (cf. Fig. 3). This confirms that it was useful to distinguish these various cases in analysis. But it also indicates that the evaluation of luminance-defined salience might be more complicated than one might have thought. It does not only require the evaluation of target-distractor contrast but also the ranking of targets and distractors and hence their differences to the background. Second, even for targets with similar contrast polarity to background and similar ranking among distractors, salience matches did not always seem to follow one common rule, but data from different test series were best fitted by different algorithms. In test series block *K*, for example, salience matches of DARK maximum targets followed the constant-addition or the constant-ratio principles (here fitted by algorithms 1, 2, and 3) but in test series block *L* were better predicted by the normalized Michelson contrast (algorithm 8). In terms of the main goal of the study, to discover simple and hopefully general rules for the computation of luminance-defined salience, these findings were discouraging and opened speculations about other yet undiscovered algorithms that might better explain the target salience in various configurations. This idea is followed up in section C below.

"Natural" rules of salience computation

In the accompanying study on luminance-defined salience of single items or blob arrays *without distractors* (Nothdurft, 2015) data have underlined two principle characteristics that reflect luminance variations of the real world and hence might be *a priori* useful models of salience computation. When the illumination of a scene changes (for example, when clouds hide the sun) the luminance of reflecting surfaces will change in a proportional way *(constant-ratio rule)*. Self-luminous targets, on the other hand, which often are brighter than other objects, will add a constant amount of luminance irrespective of illumination *(constant-addition rule)*. It would seem useful if the visual system could compensate for both these (natural) luminance variations when evaluating the salience of various items in a scene. In the previous study the luminance settings of equal-salient blob arrays were indeed found to follow the constant-ratio principle but this rule was sometimes disturbed by an additive component, in particular when targets were very bright (Nothdurft, 2015). The uncertainty of observers (also seen in some experiments of the present study) whether to apply the constant-ratio or the constant-addition (or perhaps another) rule is a frequent observation in brightness-matching experiments and apparently influenced by additional visual cues about the surface structure of the matched objects (Arend & Goldstein, 1987; Gilchrist, 1988; Robilotto & Zaidi, 2006; Schirillo, 1999a, b). Purely reflective surfaces are more likely matched according to the constant-ratio rule. In the present study, however, a strictly constant-ratio behavior (as described by the Weber Contrast, algorithm 2) was so far only seen in some test series, with DARK maximum targets in series *K, L,* and *F* (Table 2, A1-A3) and with BRIGHT maximum targets in series *L*, *LX*, and *F* (Table 2, E2 and E3). In other test series, predictions from algorithm 2 were not among the best ones, and the constant-ratio principle generally failed to optimally predict targets in minimum configurations (Table 2, rows C and G). Interestingly, the occasionally good performance of algorithm 2 with DARK maximum targets in some tests does not hold when the MSD is calculated for all test series together (Table 2, A5) but is then replaced by the constant-addition principle (algorithm 1) which had produced quite good predictions in all test series and thus wins the overall performance in tests on DARK maximum targets. For BRIGHT maximum targets, the *salmin* algorithm 5 makes the best overall predictions, closely followed by algorithm 2 (the constant-ratio principle; Table 2, row E).

In the following I will discuss three special observations in the data presented so far: (i) the on a first glance perhaps unexpected finding that the salience of DARK maximum targets did not (or only weakly) depend on background luminance (Fig. 3a and Fig. 9a); (ii) the deviations of minimum target matches in small background-distractor windows both from predictions and between dense and wide pattern configurations (Fig. 3b, *K26* and *K47*, and Fig. 10), and (iii) the peculiar bipartite fits (and partial misfits) of simple predictions to DARK targets in minimum configurations (Figs. 7a, 8a, 11 and 12).

Salience variations with background changes

If background, distractors, and targets would all represent purely reflective surfaces, then variations of scene illumination should lead to proportional luminance variations. Thus, if reference patterns and test patterns were identical and only differently illuminated, and if the visual system would compensate for such illumination variations when evaluating salience, then target adjustments should represent exactly this ratio of changed illumination (here given by a factor *r)*.

$$\begin{aligned} &bg_{reference} \rightarrow bg_{test} = r \cdot bg_{reference} \\ (3) \qquad &dis_{reference} \rightarrow dis_{test} = r \cdot dis_{reference} \\ &tg_{reference} \rightarrow tg_{test} = r \cdot tg_{reference} \end{aligned}$$

With uniform blob arrays or single items, this was indeed observed (Nothdurft, 2015). Items that followed the constant-ratio principle (as given by the Weber contrast) were seen as equally salient. With a scaling factor k that relates salience to Weber Contrast, the salience of distractors and targets on the background can be described as

$$sal_{dis:bg} \sim \frac{|dis - bg|}{bg} \quad \Rightarrow \quad sal_{dis:bg} = k \cdot \frac{|dis - bg|}{bg}$$

$$sal_{tg:bg} \sim \frac{|tg - bg|}{bg} \quad \Rightarrow \quad sal_{tg:bg} = k \cdot \frac{|tg - bg|}{bg}$$

and that of targets to distractors (Table 1, algorithm 2) as

$$sal_{tg:dis} \sim \frac{|tg - dis|}{dis} \quad \Rightarrow \quad sal_{tg:dis} = k \cdot \frac{|tg - dis|}{dis} \quad .$$

Pure illumination changes as in equations (3) would not change these results as r can be factored out and then be canceled.

In the experiments of test series blocks *K* and *F*, however, test distractors were not changed in proportion to background but were held constant. This should have changed their salience on background compared to that of reference distractors. To compensate for that also test targets had to be changed. The correction factor should be *1/r* to hold distractor luminance constant, and the same factor must be applied to test target luminance to keep the relative salience of target and distractors constant.

$$\begin{aligned} &bg_{reference} \rightarrow bg_{test} = r \cdot bg_{reference} \\ (3a) \quad &dis_{reference} \rightarrow dis_{test} = r \cdot \frac{1}{r} dis_{reference} = dis_{reference} \quad , \\ &tg_{reference} \rightarrow tg_{test} = r \cdot \frac{1}{r} tg_{reference} = tg_{reference} \, . \end{aligned}$$

In other words, the constant-ratio principle (algorithm 2) would predict that equal-salient targets *among identical distractors* must be identical, irrespective of the according background luminance.

Note however that this conclusion has two restrictions. First, it refers only to the relative salience of targets to distractors (in the following referred to as *discrimination salience)*, not to the salience of targets (and distractors) to

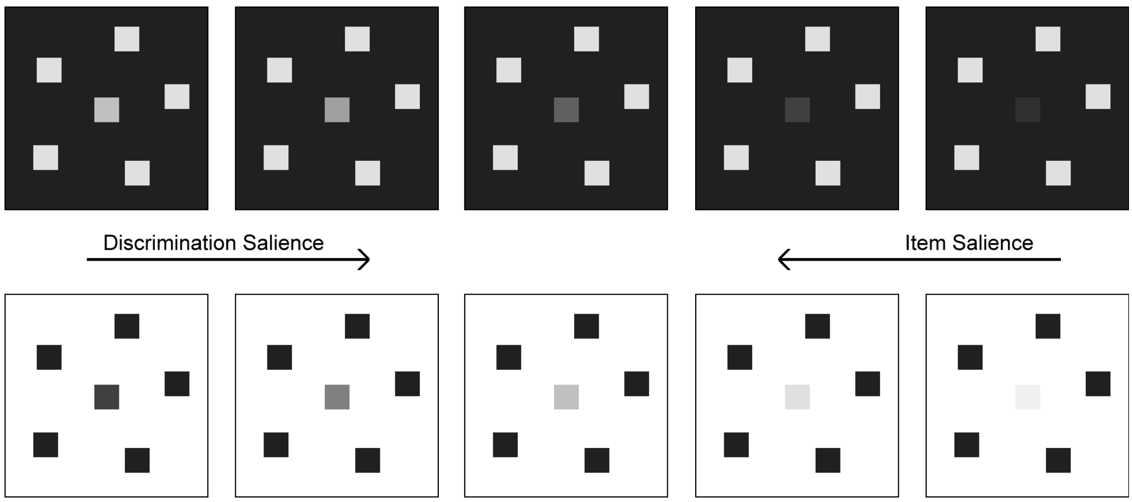


Figure 13. *Illustration of opposite salience effects in minimum target configurations.* When target luminance changes, either the discrimination or the item salience increase. In the extremes (left-hand or right-hand figures), item or discrimination salience may become very large but targets are nevertheless not salient because they are indistinguishable from distractors or invisible on the background. This opposite variation of salience effects occurs only with sparse blob arrangements and with targets in minimum configuration.

the background (in the following referred to as *item salience)*, which should indeed vary with background luminance (Nothdurft, 2015). The difference is illustrated in Figure 13. The distinction is important for understanding the deviations of minimum target matches discussed in the next paragraph.

Second, the conclusion that the constant-addition rule is not in conflict with the constant-ratio principle holds only for the special test conditions of test series *K* and *F* with identical distractors. Predictions should look different when distractor settings were not held identical (cf. equations *2)*. Proportional luminance variations of distractors and the target would then not represent constant differences. In fact, the luminance settings of equal-salient DARK maximum targets in test series block *L* were better predicted by the normalized Michelson contrast (algorithm 8), which can not be deduced from the constant-ratio principle. The normalized Michelson contrast has been reported to predict perceptual transparency in certain stimuli (Singh & Anderson, 2006) and one might have expected a much wider influence of this principle in other configurations tested in the present study. This was however not the case (Table 2).

Still, the observation that background variations do not affect the salience of DARK maximum targets but strongly affect the salience of BRIGHT targets (Fig. 9) remains puzzling. The difference should be particularly strong in dense blob arrangements while DARK targets in wide arrangements might require corrections when target-distractor differences are matched (Fig. 9a, test series F2). The differences are visualized in Figure 14 (you need a printout without contrast enhancement to see all intended variations, see Appendix) and nicely illustrate the interplay of item and discrimination salience even in maximum target configurations.

In each row of Figure 14, target and distractor luminance settings are the same and only background luminance is changed. Rows thus illustrate the results of Experiment 4 for DARK and BRIGHT maximum targets. When evaluating the *item salience* of DARK targets and distractors, you find distractors loosing their salience towards the right while all targets in the row look about equally salient. This is true for high- and low-contrasting DARK targets in wide and dense blob configurations (Fig. 14a-d). Note that this observation should be astonishing, as also the item salience of the DARK target is diminished when the background is darkened. Apparently, DARK target item salience is only little modulated by background variations if the target-to-background contrast is large. The impression is different with the BRIGHT targets, however. While high-contrasting targets (Fig. 14e, f) remain about equal-salient, BRIGHT targets with a smaller target-to-background contrast become more salient towards the right and should, for an equal-salience match, be reduced in their contrast to the background, as seen in Experiment 4 (Fig. 9c). This variation is particularly obvious in the dense blob raster (Fig. 14h). When adjusting targets for a similar *discrimination* from distractors, however, one might want to reduce target contrast from left to right (thus confirming the light-blue data curves in Fig. 9a).

Deviations in small distractor-background windows

While salience matches in dense and wide blob arrangements were often quite similar, there were notable differences in certain test conditions (Fig. 3). In these test series, matches did not only differ between the two blob densities tested but also deviated from the expected *salmin* prediction for minimum targets. Deviations were strongest in test series with small distractor-to-background differences in the test pattern (e.g., series *K43, K26* and *K47*) but nearly absent when the distractor-to-background difference was large enough (e.g., series *K25*). Similarly strong deviations were seen in Experiment 5 (Fig. 10), where looking at the data in more detail had finally helped to understand the basis of these deviations. Can we understand also the deviations seen in Figure 3?

A first hint comes from the individual matches of the three observers with test series *K47* (Fig. 15b). Only subject MCV (red symbols) had produced matches close to the *salmin* prediction, and these were similar in dense (filled symbols) and wide patterns (open symbols). In contrast, subject AJ (green symbols) had adjusted targets in the wide and the dense blob raster quite differently. The data of subject HCN (blue symbols) lay in between. The curves of *all* subjects, however, flattened with decreasing reference target luminance, and wide blob raster curves generally failed to reach the lower left corner of the framed background-distractor window.

To understand these variations it might be helpful to imagine how the stimulus patterns had looked like in these test conditions. A BRIGHT reference target (abscissa) notably darker than distractors (vertical dashed lines) but notably brighter than the background (vertical continuous

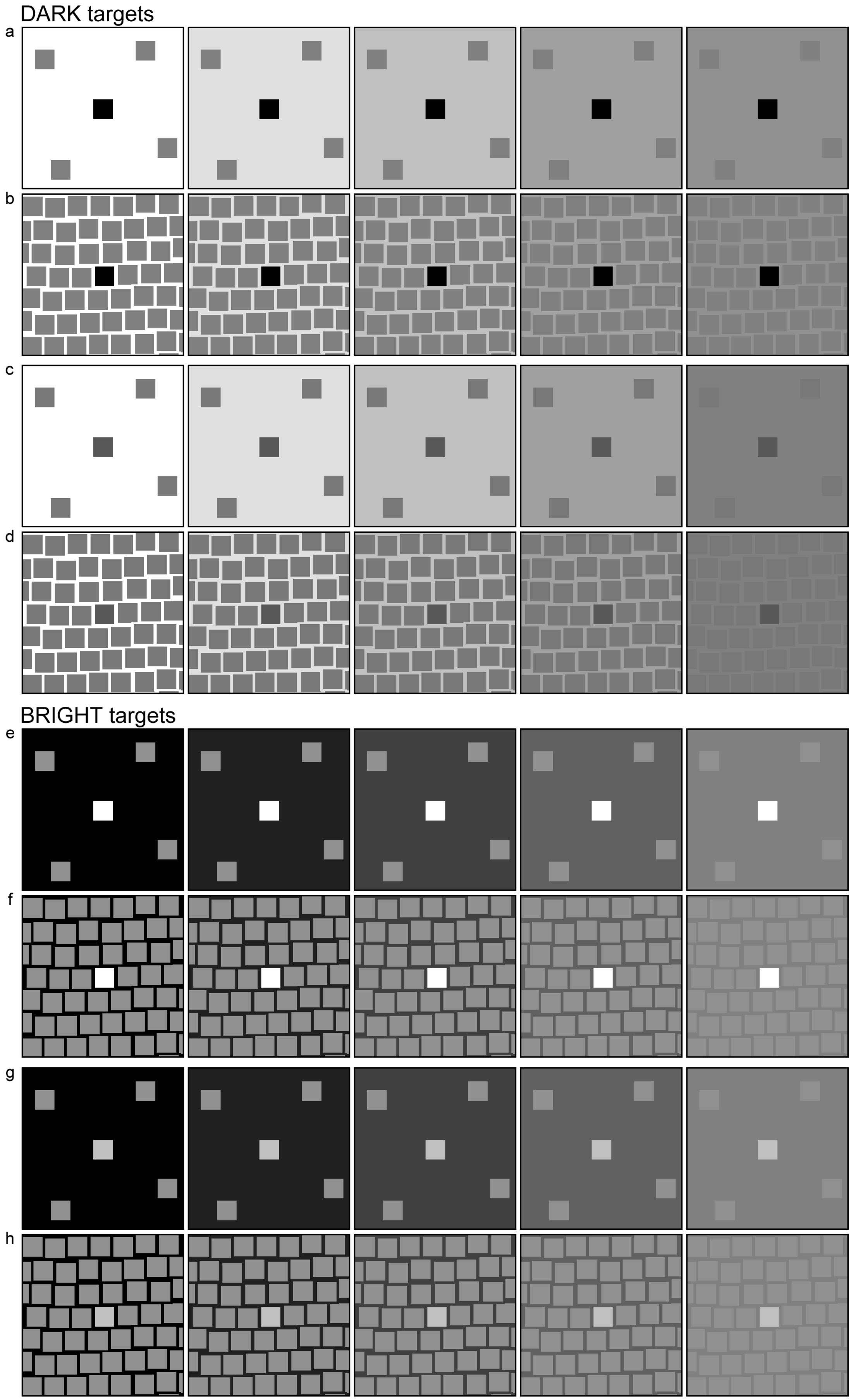


Figure 14
Legend next page

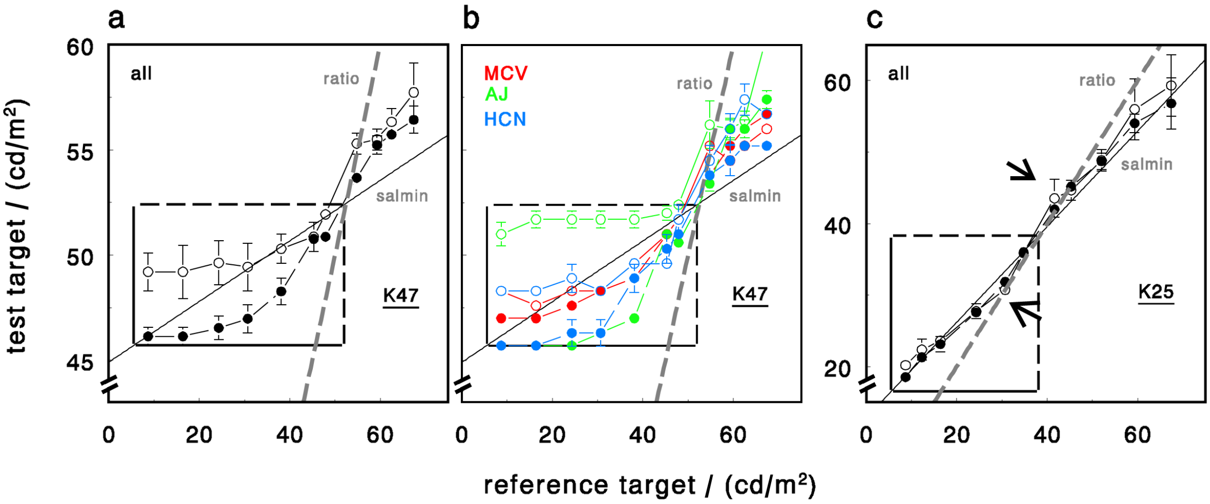


Figure 15. *Detailed look into deviations in Figure 3 (BRIGHT targets).* In certain test series of Experiment 1, there were notable deviations of equal-salience matches from the straight *salmin* predictions and between matches in wide and dense blob configurations (cf. labels 1-3 in Fig.3). The data of two test series are here re-plotted to discuss possible reasons; please note the differently enlarged *y* scales in the graphs. Curves show target-target variations as in Fig.4. Distractor and background settings were constant within each series and are shown as continuous (backgrounds) and dashed lines (distractors), respectively. Colors distinguish means and data from different subjects; matches from different blob arrangements are indicated by filled (dense) and open circles (wide). **a.** Mean data and **b.** matches of individual subjects in test series *K47* (cf. Fig.3b, label 3). **c.** Mean data with test series *K25* (Fig.3, label 1). Deviations in *(a)* and *(b)* are explained by the low item salience of test targets in wide blob arrangements. Deviations in *(c)* (arrows) reflect local shifts towards constant-ratio (also seen in Fig.10) when targets are similar to distractors. Straight lines give the salmin (algorithm 5) and the target-to background constant-ratio predictions.

lines) had to be matched with a test target (ordinate) that could differ only little from distractors and background due to the much smaller background-distractor luminance span (horizontal lines; please notice the different scales in *x* and *y*). If subjects had linearly adjusted the test target in wide blob patterns to values along the *salmin* prediction line, some targets had become barely discriminable from background. To make them as visible as the according reference targets, subjects increased the target-to-background contrast. Thus, it is not surprising that, in the wide blob raster, matches of targets approaching background luminance leveled off at a certain threshold in the small test distractor-background windows. They might have only decreased further when also the reference targets were close to background and (almost) invisible (as it was the case in Experiment 5).

But why did these deviations only occur with matches in sparsely arranged blobs (open symbols) and turned over into deviations in the opposite direction (curves steeper than predicted from the *salmin* algorithm) in dense arrangements (filled data points)? As can be visualized in Figure 2, target and distractors in dense configurations are

Figure 14 *(previous page)*. *Background effects on target salience.* **a-d.** DARK targets; **e-h.** BRIGHT targets. In each row, targets and distractors are constant and only backgrounds are varied. As a consequence, the item salience of targets and distractors diminishes from left to right. This is, however, mainly noticed for distractors which become almost invisible in the right-hand patterns. This figure illustrates the different effects from background variations as measured in Experiment 4. The salience of DARK targets in wide *(a, c)* and dense *(b, d)* blob arrangements is less affected by background variations than the salience of BRIGHT targets (*e-h)*. This is particularly obvious in *(d)* and *(h)*. While the salience of the DARK targets is about the same within the row, that of the BRIGHT targets increases from left to right. To make these targets equal-salient, their contrast must be reduced, as found in Experiment 4 (Fig.9c). Note that also the contrast of DARK targets in *(c)* should be reduced if these were not matched for equal salience but adjusted for similar discrimination from distractors. (You may need a printer with linear output characteristics to see appropriate luminance variations; see Appendix.)

arranged so close to each other that mainly the target-to-distractor contrast is seen and the target-to-background contrast becomes less important. In dense blob arrangements, therefore, there was no need to enhance the target-to-background contrast to make the target visible. Here, two subjects had instead over-enhanced the target-to-distractor contrast in the test pattern (Fig. 15b; green and blue filled circles). This too can be understood from the different distractor-background luminance windows in reference and test patterns. To match the large target-distractor contrast in the reference pattern, the test target should have quickly exceeded the much smaller target-distractor luminance span of the test pattern. Enhancements below the background level, however, were prohibited by the software in experiment; therefore deviations in dense patterns could maximally reach background luminance and thus had flattened there (cf. Fig. 15b).

Thus, the deviations in widely spaced blob patterns in test series *K47* (Fig. 3, label 3) are likely the results of two different salience effects, the *discrimination salience* that lets the target stand out from distractors, and the *item salience* that lets the target stand out from background. Although experiments were designed to measure (and match) discrimination salience, i.e. the strength at which the target differed from distractors, subjects apparently made additional adjustments when item salience was too small. This effect was only seen in wide blob configurations and only with targets in minimum configuration, since only these could have produced (too) faint targets among stronger distractors. For targets in maximum configurations item salience was always larger than discrimination salience.

This interpretation is supported by the findings in Experiment 5 (Fig. 10). The finer resolution of luminance variations in that experiment and the inclusion of reference targets close to background had obviously helped the subjects to overcome threshold settings in luminance adjustments. But the fact that almost all subjects adjusted background-near targets according to the constant-ratio principle indicates that matches were made for equal item salience (target-to-background differences) rather than equal discrimination salience (target-to-distractor differences).

Deviations as in Experiment 5 were also seen with large background-distractor windows in Experiment 1 when target luminance was close to *distractor* luminance (e.g., *K25* in Fig. 3; label 1). Data points of matches are locally shifted away from the *salmin* predictions towards the identity line (here representing the constant-addition and the constant-ratio rules). These (relatively small) local deviations were seen with both dense and wide blob configurations and with both maximum and minimum targets (Fig. 15c, arrows).

Multiple or bipartite fits

One peculiarity in the fits was the matching performance in test series block *O*. The original data were closely predicted by algorithm 10 (Fig. 8a, left-hand graph), but could alternatively be also predicted by a bipartite fit of the *salmin* algorithm 5 (when test distractors were brighter than reference distractors) and the constant-addition algorithm 1 (when test distractors were darker than reference distractors; Fig. 7a). In a later follow-up experiment (Exp. 6) however, neither algorithm 10 nor the bipartite fit of algorithms 5 and 1 could reliably predict the data of all test series (Fig. 11 and 12). Instead, the experiment showed that matching performance was rather variable and perhaps even uncertain for some test conditions.

To see if such an uncertainty might have also been present in the matches of Experiment 3, the original data are re-analyzed and re-plotted in Figure 16. In addition to the means (re-plotted in Fig. 16a) also the individual matches are shown (Fig. 16b). As is particularly obvious with test series *O12,* response variations are similar to those in Figures 11 and 12. When test distractors were brighter than reference distractors (data points on the left), the matches strayed little and mainly around the *salmin* prediction. But when test distractors became darker than reference distractors (data points on the right), the individual matches were distributed more widely, in particular in the transition conditions when test distractors were only little darker than reference distractors *(O11)*. Test targets were then sometimes matched even equally bright to the reference target. In none of the curves was there an obvious “switch” from one algorithm to the other. The uncertainty at intermediate settings has produced mean data that are, in this particular experiment, best described by algorithm 10. But Experiment 6 has shown that the same matches in other samples of test conditions may differ and are then not adequately fitted by this algorithm.

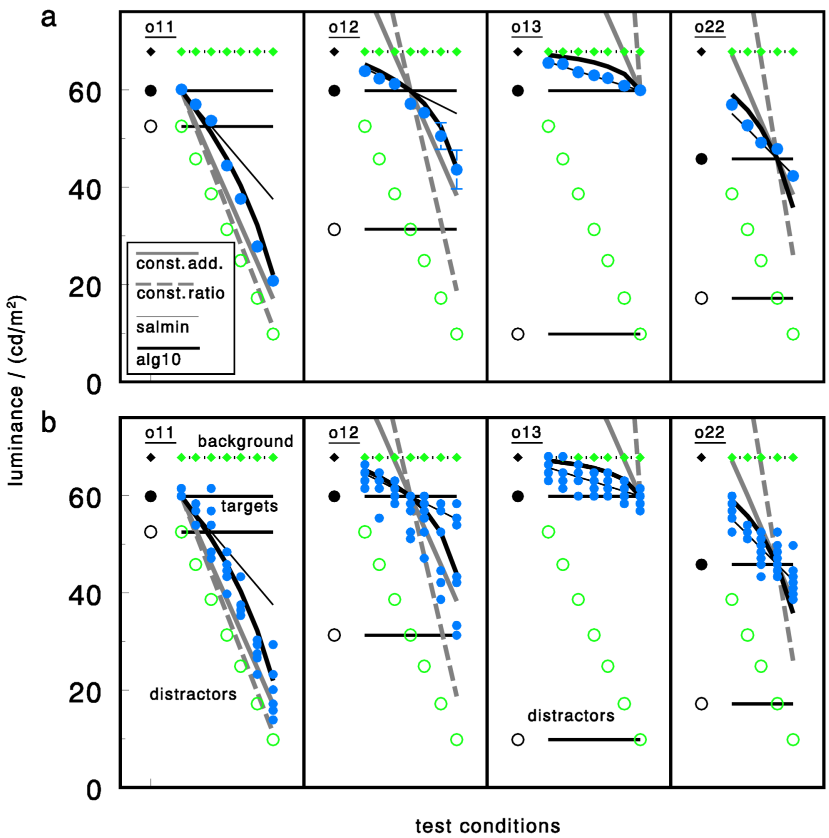


Figure 16. *Re-plot of the wide blob raster data in Figure 7.* **a.** Means of wide-raster matches and **b.** according single trial matches. Experiment 6 has suggested different matching rules for different target conditions, which led to an increased response variability for intermediate settings. This re-plot was made to look for similar variations in the data of Experiment 3. Indeed, matches varied around the *salmin* prediction (thin black lines) when test distractors were brighter than reference distractors (left-hand sides in each graph and entire series *O13)* and tended to approached constant-addition or constant-ratio rules (gray lines) when test distractors were much dimmer than reference distractors (far right-hand sides of graphs). In intermediate ranges matching variability was increased. In some (but not all) test series, the best fit of the resulting data was made by algorithm 10 (thick black lines).

Conclusions from Discussion:

Demos illustrate the difference between *item salience* and *discrimination salience* (Fig. 13) and the different sensitivity of DARK and BRIGHT targets to background variations (Fig. 14). – Deviations from simple predictions in Figure 3 can be explained by large differences between background-distractor settings in reference or test patterns. – Uncertain transitions between salience computation rules may lead to singular and seemingly peculiar fits.

Summary of Section A

From a large number of test series in which two targets were matched for salience, we can now draw the following conclusions.

1. Matching performance appeared to follow systematic rules, that is, subjects made similar adjustments, and target settings varied in an apparently regular way, when various luminance parameters of the patterns were changed.

2. However, there was no single, common rule of how salience of different targets is computed, but salience computation depended on various stimulus parameters such as target luminance polarity to background, target-to-distractor ranking and, in some conditions, also the density of target and distractor arrangements.

3. Within each of these target groups luminance settings of equal-salient targets can often be predicted, but formulas may require more distinctions than were initially made.

4. In addition to all these variations, target salience (even just in the luminance dimension) is defined from several properties, and different salience effects may add or cancel each other. Targets are salient because they differ from distractors *(discrimination salience)*, but also because they differ from background *(item salience)*.

5. Subjects, by attending to different such salience components, could identify the same pair of targets as exactly equally salient or not. This might create a general

problem for salience matches as in the present study. Despite this variability, however, there was a large consistency in the subjects' performances, and even uncertainties in matching decisions seemed to be averaged out when several matches were repeated.

B: SALIENCE MATCHES OF DIFFERENT TARGETS

Hints for reading: This section reports salience matches of *different* target types. Experiments 7-10 compare targets in maximum and minimum configurations (look it up in Fig.1) with partly a replication of maximum to maximum target matches from section A. The new aspect here is that test patterns serve as a constant measure of target salience. Experiments 11 and 12 compare bright and dark targets in various configurations. Experiment 13 compares the salience of targets in the maximum-minimum paradigm.

All experiments so far have compared same target types in similar conditions, that is, reference and test targets were either both dark or both bright, and were both presented in either the maximum or the minimum configuration. These restrictions are given up in the following experiments.

Experiments 7-10: Matches of similar targets in maximum and minimum configuration *(test series block J)*

In most experiments of section A, a constant reference pattern had been compared with various test pattern settings to find the systematic luminance variations underlying equal salience. The opposite procedure was followed in Experiments 7-10. Various reference patterns were compared with one common test pattern, in which background and distractor settings were held constant and only target luminance could be adjusted. This had several advantages. The major advantage was that the constant test pattern could now serve as a constant measure to compare the target salience of different reference patterns, which had not been possible in the previous experiments.

One goal of this new approach was to compare the salience of targets in different presentations, like targets in maximum and targets in minimum configuration. Experiment 1 has shown that the algorithms to compute target salience in these two conditions might be different (cf. Fig. 3a); it would therefore be helpful to compare the salience of these targets to a constant measure. The previous experiments have further suggested that targets in the maximum configuration are easier to adjust than targets in the minimum configuration, which are affected by item and discrimination salience; therefore, test patterns with dark or bright *maximum* targets were chosen as the comparison measure. To allow for a maximal range of possible target adjustments, backgrounds were set to the monitor limits (low for bright targets; high for dark targets) and distractors slightly above or below so that they were well seen but did not too much restrict the remaining range for target adjustments.

In addition to the intended comparison of minimum and maximum target salience, the new procedure was also used to repeat, and hopefully confirm, some of the previous matches of targets in maximum conditions. The adjustments of a fixed salience meter should now allow us to compare the salience strength of different such patterns (which had not been possible with the experiments of section A). Altogether, the various target combinations created four experiments with different blocks of test series.

Stimuli

Twenty test series were created for Experiments 7-10 and grouped in four blocks associated with the four experiments. Two blocks (Experiments 7 and 9) displayed DARK targets, and two blocks (Experiments 8 and 10) BRIGHT targets. DARK targets were matched with DARK targets and BRIGHT targets with BRIGHT targets. In one of each two blocks, reference targets were shown in maximum configuration (Experiments 7 and 8), in the other two blocks in minimum configuration (Experiments 9 and 10); test targets were always presented in maximum configuration.

Every block contained six or seven test series that were presented in five experimental runs with 8-9 stimulus patterns each. Background luminance was constant over each experiment and was always the same in reference and test patterns. Distractor and target luminance settings in reference patterns were systematically varied over the different test series, as illustrated in Figure 17, for example. Distractor settings in the test patterns were constant and only targets could be adjusted. In each single run of a test series, all stimulus conditions were presented twice, in pseudo-random sequence, with exchanged locations on the screen. The observers' task was always to match the two targets for salience, as in all previous experiments of the study.

The same three subjects that had performed Experiments 1-4 also performed Experiments 7-10. BRIGHT targets in wide blob configurations were sometimes hard to match, and these tests (Experiments 9 and 10) were not performed by all subjects.

To estimate the influence of certain parameters in the matches of Experiment 9 (DARK minimum and maximum targets), three additional test series were performed by subject HCN, in which either the background in the reference patterns or the distractor level in the test patterns were set to new values. These test series included all test conditions of Experiment 9 (which did not fall outside the new parameter settings) and were only run on dense blob raster conditions.

Results and Discussion

<u>Targets in maximum configuration</u> (Experiments 7 and 8)

The matches of reference targets in maximum target configuration are partly confirmatory; similar matches had already been made in test series blocks *K, L,* and *F* of section A. However, in none of the previous test series could the target salience itself be measured.

Figure 17 gives an overview of the matches with *DARK maximum targets* (Experiment 7). Distractor and target settings in the reference patterns were systematically varied (Fig. 17a); the various test conditions in the figure are aligned so that reference targets fall upon straight lines (gray). The according matches in the test patterns (Fig. 17b) show a similar target variation but now touched to the different (and, in fact, constant) distractor luminance

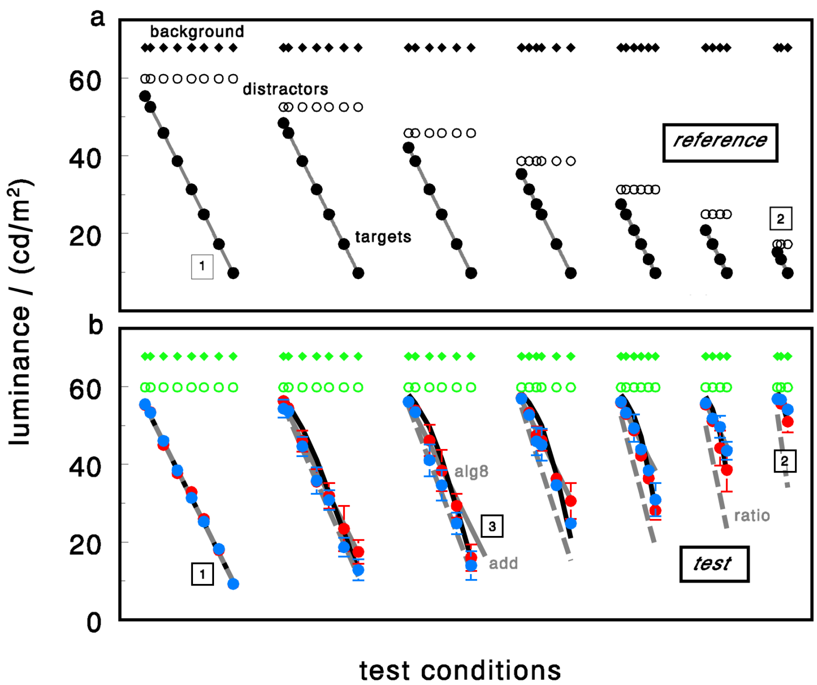


Figure 17. *Equal-salience matches of DARK maximum targets in Experiment 7.* The new step in this and the following experiments is that matches are made with a constant test pattern setting (in which only test targets were adjusted), which could thus serve as a constant meter of reference target salience. The figure shows all test conditions of Experiment 7. Symbols as in previous figures. **a.** Luminance settings of reference patterns; on constant background, distractor and target luminance settings were systematically varied. **b.** According matches of test targets on the same backgrounds, with constant distractor settings. Data points were shifted in *x* to let reference targets fall upon straight lines; the same shifts were used in the presentation of test pattern data. Matches were obtained with dense (red circles) and wide blob arrangements (blue circles). Predictions of the data from various computational algorithms are shown by superimposed lines. Best fits were obtained with the constant-addition rule and with algorithm 8. Framed numbers serve as labels referred to in text.

setting in the test patterns. We know (Nothdurft, 2015 and unpublished data) that the salience of DARK targets on constant background varies with the luminance contrast of targets and background. Thus, the test targets at label 1 (Fig. 17b) were more salient than the test targets at label 2. From comparison with the corresponding reference targets in Figure 17a we can therefore conclude that DARK targets among very dark distractors (right-hand test series in Fig. 17a; label 2) are less salient than the same DARK targets among less dark distractors (left-hand test series, label 1). Interestingly, test targets at intermediate curves (e.g., label 3) were sometimes adjusted to similarly large amplitudes as the test targets at label 1, although the according reference distractors were already dimmed.

Of particular interest is whether the target-to-distractor ("discrimination") salience would relate to constant addition, constant ratio (Weber contrast) or other simple computation rules. In series *K* and *F* of section A (Experiments 1 and 4), matches had followed the constant-addition and constant-ratio rules, which could not be distinguished in these tests. In series *L* (Experiment 2), however, where these algorithms could be distinguished, matches were best fitted by predictions from algorithm 8. In the present data (Fig. 17), the constant-addition principle (|*tg-dis*|=*constant)* would let test targets fall upon lines parallel to those of the reference targets (Fig. 17b, continuous gray lines), whereas the constant-ratio principle *(tg/dis=constant)* would cause them to cover the full luminance range of distractors and background (dashed gray lines). The data apparently show both; the similarity of reference and test target slopes at labels 1 and 2 supports the constant-addition model, but the similar range of target settings at labels 1 and 3 the constant-ratio model. Constant ratio, however, does not predict the matches of targets among dim distractors (curves at label 2 and the three curves to the left of it). In fact, the various matches are rather well predicted by algorithm 8 (thick black lines) which produces similar MSD values as the constant-addition algorithm 1 (cf. Table 2, A4).

The different predictions from the constant-addition and constant-ratio principles in Experiment 7 are illustrated in the target-to-target scatter plots of Figure 18. While the constant-addition principle predicts that data points fall upon parallel lines with slopes of $m=1$ (Fig. 18a), the constant-ratio principle would predict that they should fall upon lines with different slopes ($m\geq1$) all pointing to the origin (Fig. 18b). The experimental data (Fig. 19) do not show such a systematic slope variation, neither with dense (Fig. 19a) nor with wide blob configurations (Fig. 19b). Linear regression lines of the data reveal slopes around $m=1$, with no systematic shifts. And most data points fall closely upon straight lines, not on the curved lines that should be expected from algorithm 8 (Fig. 18c).

Conclusions from Experiment 7:	
Target type	*Salience variations follow*
DARK maximum	constant-addition rule (algorithm 1)

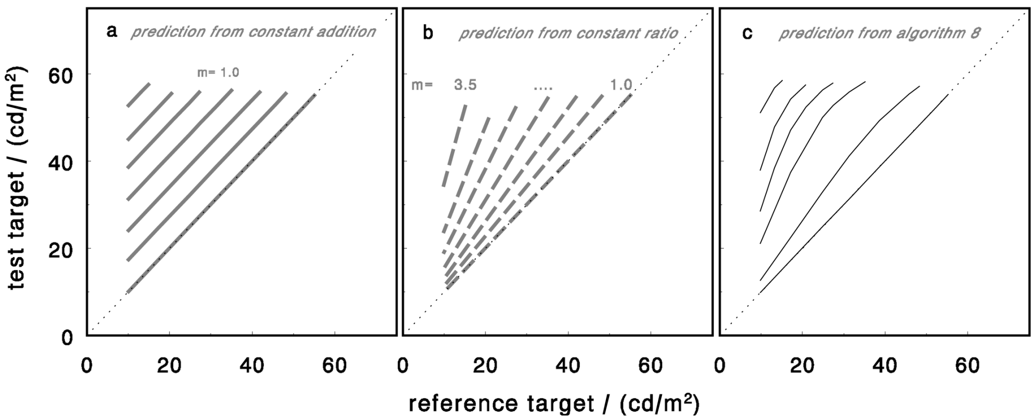


Figure 18. *Predictions of target-to-target variations in Experiment 7.* Curves show the expected courses of target-to-target variations when salience were based on certain salience computations (Table 1); **a.** constant addition (algorithm 1); **b.** constant ratio (algorithm 2); **c.** algorithm 8. Comparison with Fig.19 reveals *(a)* as the most likely basis of salience matches in this experiment.

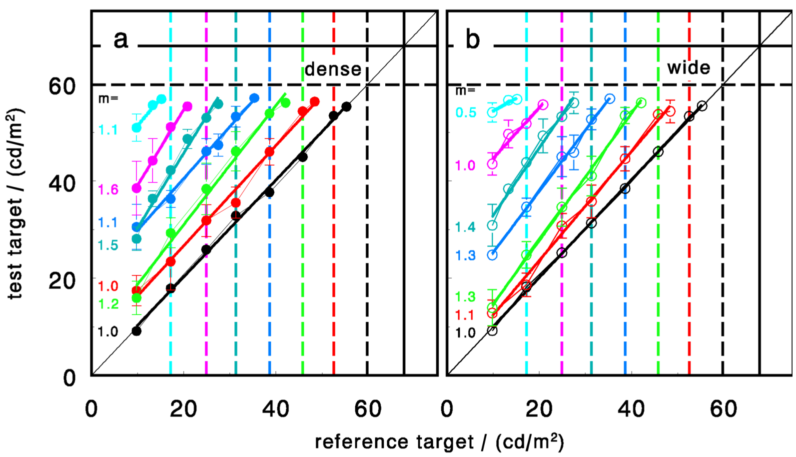


Figure 19. *Measured target-to-target variations of DARK maximum target matches in Experiment 7.* **a.** Dense blob arrangements; **b.** wide blob arrangements. Continuous black vertical and horizontal lines indicate background luminance, which was identical in reference and test patterns in all tests of Experiment 7. Dashed horizontal and vertical lines indicate the distractor settings which also were identical in all test patterns (horizontal black line) but varied across reference patterns (color-coded vertical lines); the according data points are plotted in the same color. Thick oblique lines are regression lines through same-color data points; their slopes are listed aside. Data points and regression lines tend to run in parallel to the identity line (test target = reference target) and thus support the constant-addition model.

An analogous overview of the *BRIGHT maximum target matches* (Experiment 8) is shown in Figure 20. Again, distractors and targets were systematically varied in the reference patterns (Fig. 20a) but only targets were adjusted in the according test patterns (Fig. 20b, c). Backgrounds were same and constant. BRIGHT target matches in the wide blob raster turned out to be difficult and variable, and were performed by only one subject. For analysis, the data from dense and wide blob patterns were therefore pooled. Predictions from constant ratio (gray) and the *salmin* algorithm (black) both seem to fit the data quite well except for the left-hand curve where test distractors were brighter than reference distractors and test targets should have been adjusted to much brighter luminance settings if the predictions were to be fulfilled.

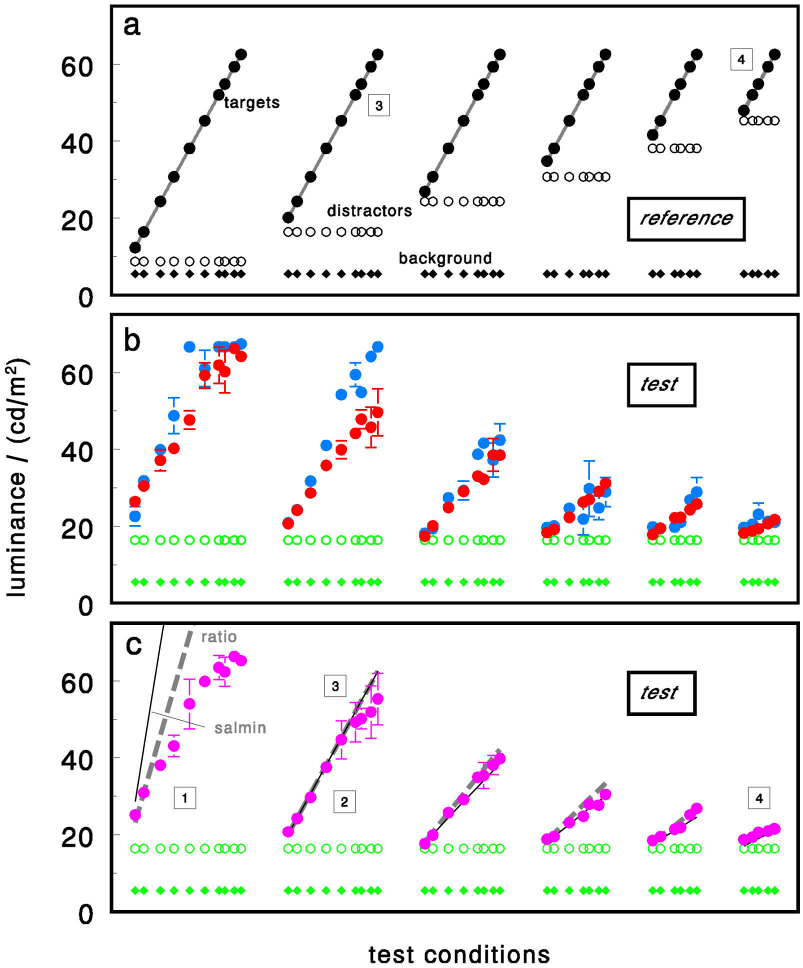
In the experiments of section A, BRIGHT targets followed, beside algorithm 8, either the *salmin* algorithm (series *K* and *F;* Table 2, E1, E3) or the constant-ratio principle (series *L* and *LX*; Fig. 5b, c; Table 2, E2). Both algorithms made similar but not identical predictions in Experiment 8. They only differ in the divisor (*sal~|tg-dis|/dis*, for constant ratio; *sal~|tg-dis|/|bg-dis|*, for the *salmin* algorithm; see Table 1). This difference can be visualized with the scatter plots in Figure 21. All curves have to cross the intersection of the according distractor levels (dashed lines). Prediction curves from the *salmin* algorithm will also cross the intersection of background levels (as shown in Fig. 21b), whereas prediction curves from the constant-ratio principle would intersect at the origin (zero luminance; not shown). With the low background settings of Experiment 8, these differences are small.

Both predictions closely fit the data of all but one curves (Fig. 21a; cf. Fig. 20c). The left-most curve (black in Fig. 21a and steepest prediction line in Fig. 21b) with test targets brighter than reference targets is not met by these predictions. Instead, the data curve runs parallel to the identity line (dotted in Fig. 21a) thus following the *constant-addition* principle. Deviations are only seen when test targets should have exceeded the limits of the monitor (gray-filled data points near label 1). This confirms earlier observations (cf. Experiment 2, and Nothdurft, 2015) that targets which are the brightest items in a scene may be frequently adjusted according to the constant-addition rule.

The close fit of the data in Figure 20c is reflected in small MSD values (Table 2, E4). The MSD values for the *salmin* algorithm 5 and the constant-ratio principle (algorithm 2) are similar and both small, when the left-most, deviating curve (label 1 in Fig. 20c) is not included.

Note that the test series also included identity matches (red data in Fig. 21a, labels 2 and 3 in Fig. 20c), in which reference and test patterns displayed identical background and distractors settings. Perfect adjustments should have revealed identical target settings in reference and test patterns. This was, however, not exactly found but matches of very bright targets deviated from identity (label 3 in Fig. 20c). The deviation might have been caused by two effects, (a) the restricted exploration range when target luminance was close to the monitor limit (which might have biased subjects to adjust targets dimmer than in a perfect match), and (b) the decreased sensitivity to luminance variations in bright targets (Nothdurft, 2015). But neither effect could explain the

Figure 20. *Equal-salience matches of BRIGHT maximum targets in Experiment 8.* Symbols and presentation as in Fig.17. **a.** Luminance settings of reference patterns; **b.** luminance settings of salience-matched test patterns; **c.** averaged data from dense and wide blob patterns in *(b),* with predictions from constant ratio (dashed gray lines) and the *salmin* algorithm (thin black lines). Both predictions are similar and, except for the left-hand data curve (label 1), closely fit the data. The matches at labels 2 are identity matches (same backgrounds, same distractors) which however deviated at large luminance settings (label 3). Same targets among different distractors varied notably in salience (cf. labels 3 and 4).



systematic switch from salience computation along constant-ratio or *salmin* algorithms in the right-hand curves to salience computation along constant addition in the left-most curve 1.

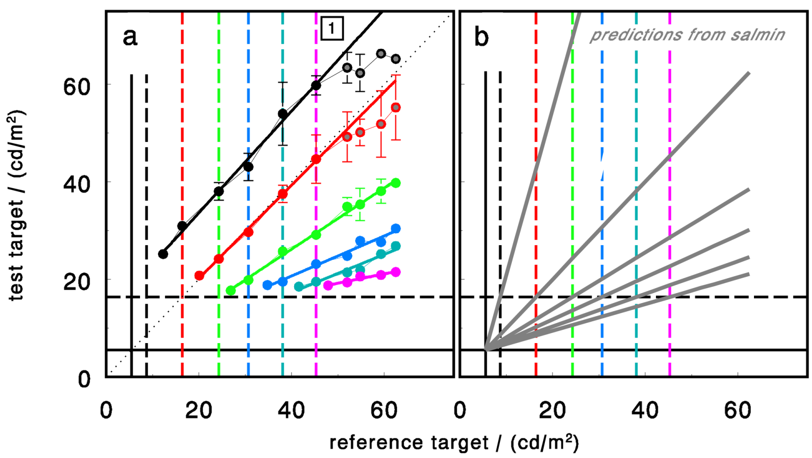


Figure 21. *Scatter-plots of target-to-target variations in Experiment 8 (BRIGHT maximum targets).* Presentation as in Fig.19. **a.** Matches from Fig.20c. Vertical and horizontal lines indicate background (continuous) and distractor settings (dashed). Thick oblique lines are regression lines through the data points (except those with gray kernels). **b.** Predictions from the *salmin* algorithm 5 (cf. Table 1). Predictions from constant ratio look very similar except that lines would cross the origin. The label number is referred to in the text.

We can thus conclude that the salience of BRIGHT maximum targets follows the *salmin* or the constant-ratio algorithm as long the luminance range of the reference pattern is not exceeded, but that there is a tendency to weaken occasional bright targets outside this range and match them according to the constant-addition rule instead. This corresponds to observations made with homogeneous blob arrays (Nothdurft, 2015).

Note that the luminance of test targets in Figure 20c is directly (but not necessarily linearly) related to their salience. We can thus draw similar conclusions about the salience of identical targets among different distractors, as we did with DARK targets before. The salience of a BRIGHT maximum target among bright distractors is reduced when the luminance difference between target and distractors is diminished (cf. targets at labels 3 and 4 in Fig. 20a and c). Different to DARK targets, however, this reduction is not linearly related to the luminance difference between target and distractors but is normalized to the (simultaneously increasing) distractor-to-background luminance change. That is, BRIGHT maximum targets loose their salience faster than DARK targets when the item salience of distractors is increased.

Conclusions from Experiment 8:	
Target type	*Salience variations follow*
BRIGHT maximum	salmin (algorithm 5) or constant-ratio (algorithm 2)

Targets in minimum configuration (Experiments 9 and 10)

Minimum targets were so far not compared with maximum targets and the outcome of Experiments 9 and 10 cannot be easily predicted from the experiments in section A. At least for DARK targets, the algorithms predicting matching performance with either target type were different (cf. Table 2, A5 vs. C5). This leaves room for new and unexpected observations, which then need to be evaluated and interpreted.

Indeed, when matching *DARK targets* (Experiment 9) in wide blob configurations, a first surprise occurred. Figure 22 shows the various reference pattern settings (black) together with the matched test pattern settings of the individual observers (colored). Note that the discrimination salience of the DARK minimum reference targets increases when target luminance is *increased* (and targets become more distinct from distractors) but that of the DARK maximum test targets increases when target luminance is *decreased* (and targets also become more distinct from distractors). Thus, on a first view the data in Figure 22 show that an increasing (minimum) target salience in reference patterns is accompanied by an increasing (maximum) target salience in test patterns, as one should expect. The figure also shows that in reference patterns with a small luminance range (right-hand curves) target salience is generally reduced compared to patterns with a larger luminance range between background and distractors (left-hand curves). The general outline of salience variations with minimum targets is thus not different to what had been found with maximum targets. A surprise however, was the matching of targets in wide blob configurations. While subject AJ made similar adjustments for targets in dense (red) and wide blob configurations (blue), subject HCN performed quite differently in the two tasks. Particularly interesting are the data of subject MCV, who had performed differently in two runs on the wide blob raster. In the first run (dark-blue data points), her matches were similar to those obtained with dense blob configurations (and, by and large, similar to those of subject AJ, although at different sensitivity). In a later run (light-blue data points), however, her matches were quite different and for brighter reference targets much closer to the distractors (indicating reduced target salience). From the discussion in section A we must assume that in the second run the subject did not only evaluate the differences between targets and distractors but also the target's *item salience* (based on the difference of targets to background). Minimum targets near background luminance are strongly different from distractors but are simultaneously very similar to background, and hence display little or no item salience (cf. Fig. 13). Apparently, subject MCV had switched from evaluating solely target-distractor differences in the first run to evaluating true target salience in the second run, whereas subject AJ had matched the target-distractor differences in all her runs. Subject HCN matched salience (in its combination of discrimination and item salience) in all his runs right from the beginning. Similar variations in performance were not seen in dense blob configurations, as item salience occurs predominantly in wide but not dense blob arrangements. Note that, although the preference of subject AJ for discrimination contrast might have been switched by special instructions, the rivaling effect of item salience was nevertheless present and led to an increased variability of target matches in some of her adjustments.

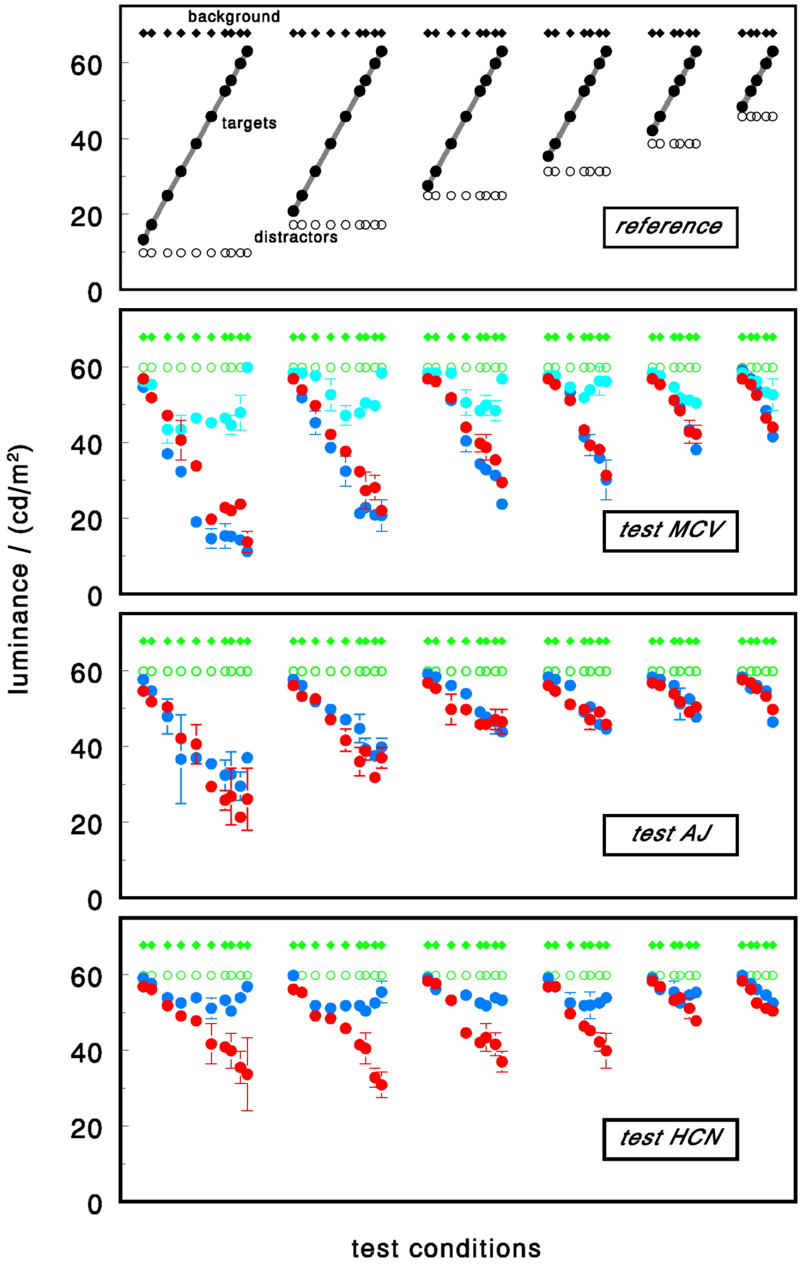


Figure 22. *Equal-salience matches of DARK minimum and maximum targets (Experiment 9).* Symbols and presentation of luminance levels as in Fig.17; reference pattern settings in black, matched test pattern settings in color. The individual matches of three subjects show different response characteristics in wide blob patterns (blue data points), which did (HCN) or did not (AJ) take care of item salience. Subject MCV had switched from one mode to the other in different runs (different blues).

Because of the different performances, the data from wide blob configurations were split and further analyzed in separate samples (Fig. 23). One data sample (*"wide sample1"*; dark-blue symbols) included the matches of subject AJ and of the first run of subject MCV; the other sample (*"wide sample2"*; light-blue symbols) included the matches of subject HCN and of the second run of MCV. Matches with *dense* blob configurations were averaged from all three subjects (red symbols). Note that the smaller sample and the different sensitivity of subjects MCV and AJ (cf. Fig. 22) have strongly increased the s.e.m. of the mean data in the *wide sample1* data*;* these variations were smaller in the *wide sample2* (less data variation) and the *dense* configuration data (three subjects).

Scatter plots of salience-matched targets in dense and wide configurations are shown in Figure 24. With dense blob configurations (Fig. 24a) data points fall upon parallel straight lines with slopes of m=*-0.53* to m=*-0.68*

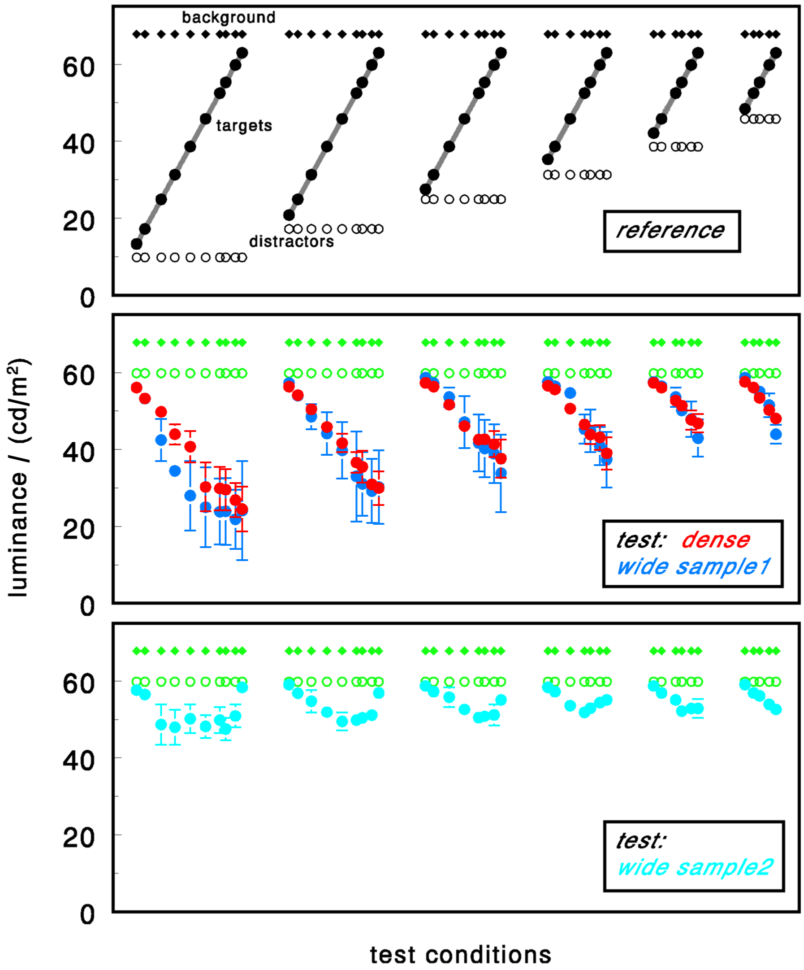


Figure 23. *Averages of data in Figure 22.* The different response characteristics in Experiment 9 required different averages of the data; matches in dense blob patterns were averaged from all subjects (red circles); matches in wide blob patterns were split into matches that had ignored item salience ("sample1"; AJ and first run of MCV; dark blue circles) and matches that did reflect the vanishing item salience ("sample2"; HCN and second run of MCV; light-blue circles).

(mean slope m=*-0.61* ± *0.05*). The same is true in the wide sample1 data (Fig. 24b), although slopes are generally steeper (mean m=*-0.73* ± *0.07)* and some curves are clearly bent rather than straight (e.g., black data points). The slope of the black curve is quite different if the regression line is fitted to only the first five data points (m=*-1.16;* dashed regression line). However, the wide sample2 data look quite different (Fig. 24c). Curves start off from low target-to-distractor contrast (colored dashed lines indicate the according distractor settings) at slopes similar to those in dense patterns, but then quickly reach a plateau while the reference target-to-distractor contrast is further increased, and finally drop back to values near test distractors (low test target salience) at maximal target-to-distractor and minimal target-to background contrast in the reference patterns. This "drop-back" reflects the reduced item salience of reference targets when they approach background luminance (continuous vertical black line) and was seen in all curves at the two highest reference target settings tested. Thus, item salience predominated the matches mainly at these two target luminance settings (63.0 cd/m^2 and 59.8 cd/m^2, on 67.8 cd/m^2 background luminance) and was less important from the third-highest tested target setting on (55.3 cd/m^2 and below).

It is interesting to look at algorithms that could predict these results. While the equal-salience matches of DARK maximum targets in section A (and in Experiment 7) were well predicted by algorithm 1 (sometimes indistinguishable from algorithms 2 and 3) and algorithm 8 (Table 2, row A), the best predictions of DARK minimum target matches were obtained from algorithms 10, A, and others, but not algorithms 1, 2, or 8 (Table 2, row C). In the scatter plots of Figure 24, combinations of algorithm 1 should have produced straight and parallel lines at a slope of m=*-1.0* (as the salience of the two targets increases with increasing and decreasing luminance, respectively).

Straight lines are indeed found in the data, but at a different slope (mean slope *m=-0.61).* Thus, for equal salience, luminance variations of the minimum target were accompanied by about 60% smaller luminance variations (in the opposite direction) of the maximum target. This constant relationship of minimum and maximum target variations in the dense blob raster is an interesting observation that may require further evaluation in a separate study. Test target modulation was strongly reduced in the *sample2* data from wide blob configurations.

The constant slopes in Figure 24a had suggested three modifications which were followed up in later test series by subject HCN. In one modification, the distractor level in the test patterns was changed (Fig. 24d), in the other two modifications the background of reference patterns was diminished (Fig. 24e and f). All test conditions outside the remaining test ranges were removed. Tests were performed only on patterns with the dense blob raster. Even under these modifications, data curves were about parallel with similar slopes as in Figure 24a.

Note however that the constant variations of DARK minimum and maximum targets (parallel lines in Fig. 24a and b) are unexpected. In Experiment 1, equal-salient DARK minimum targets were scaled to the according background-distractor luminance span (*salmin* algorithm 5) which has led to curves with different slopes in Figure 3. Thus one should also expect that each curve in Figure 24a would activate the same full salience variation and curves would hence increase in steepness from the left to the right instead of running nearly in parallel. Data from additional tests, however, confirmed both reported findings. They showed nearly parallel data curves in various DARK minimum to DARK maximum target

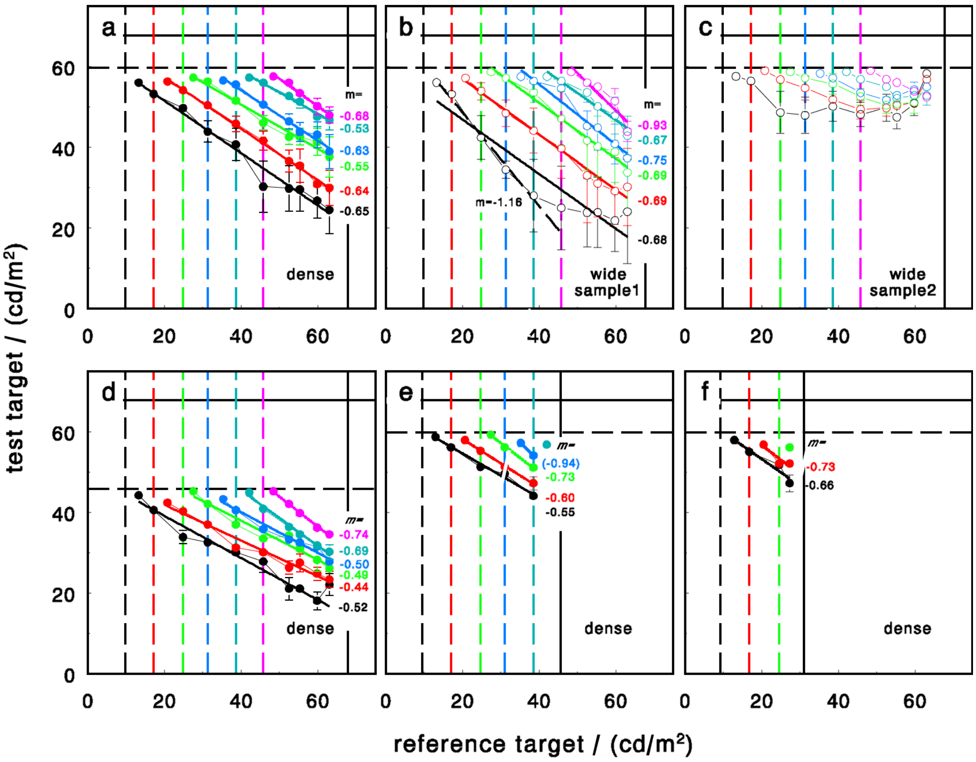


Figure 24. *Target-to-target variations in Experiment 9*. As in the scatter plots before, colors distinguish test series, symbols the data from wide (open circles) and dense blob patterns (filled circles). **a.** Matches in the dense blob raster; **b.** matches in the wide blob raster, data sample1; **c.** matches in the wide blob raster, data sample2. All these data are taken from Fig.23. Vertical and horizontal lines indicate background (continuous) and distractor settings (dashed), and thick oblique lines are regression lines through the data points. **d-f.** Additional matches from subject HCN with different background and distractor settings. All matches are represented on nearly parallel lines with slopes of $m \approx -0.6$ indicating that luminance variations in DARK minimum targets were equal-salient to about 60% luminance variation in the DARK maximum targets.

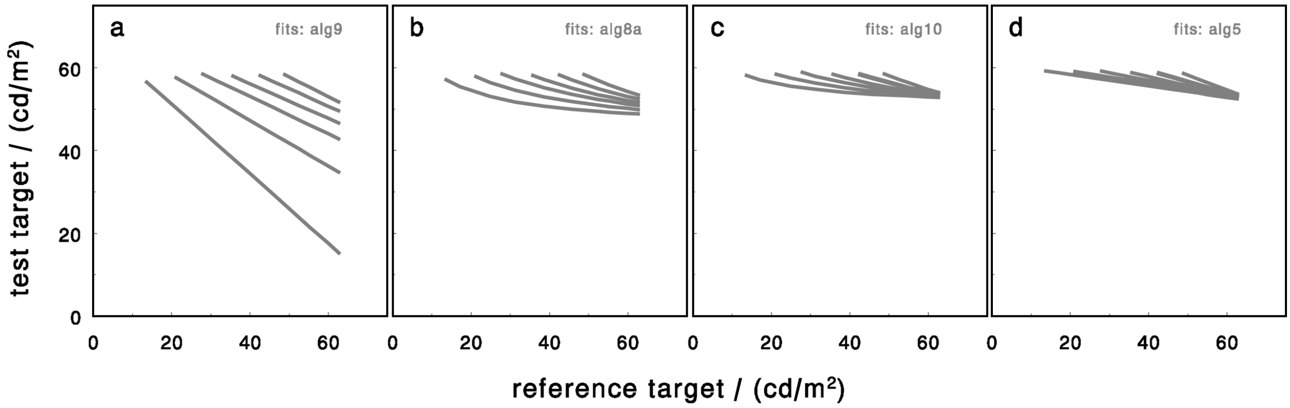


Figure 25. *Target-to-target variations in Experiment 9 as predicted by different salience algorithms (Table 1).* Curves plot the variations to be seen in Fig.24 when predicted by "unique" algorithms (see text); **a.** algorithm 9; **b.** algorithm 8a; **c.** algorithm 10; **d.** algorithm 5. These predictions provided the best fits to the data but do not resemble all response characteristics seen in Fig.24.

matches (as in Experiment 9; cf. Fig. 24d-f) and also scaled matches of DARK minimum targets against each other for the according background-distractor luminance span (as in Experiment 1). An explanation of this seeming inconsistency can not yet be given.

Can we predict the matches in Figure 24 with simple salience algorithms? The comparison of two targets in different configurations opens, in principle, two ways for computational predictions. We may select the best algorithms for each target type (for DARK targets, for example, algorithm 10 for minimum and algorithm 1 for maximum configuration; cf. Table 2 column 5) and combine them to predict the current matches *("mixed combinations")*. Or we may assume that salience matches in a given stimulus are achieved from a common process with only one algorithm (for example the constant-ratio rule) being applied to both targets *("unique algorithms")*. The difference of both procedures is explored in section C. Here, we will look at certain straight-forward predictions and test whether they can explain the data. Neither the constant-addition (algorithm 1) nor the constant-ratio rules (algorithm 2), which both had predicted several matches in section A, could even weakly predict the present data. Table 3, A1/2 and Table 3a, row A list the algorithms with the smallest MSD values, and neither algorithm 1 nor algorithm 2 are listed there. An exception is the *salmin* algorithm 5, which produced fairly good MSD values with the *wide sample2* data (Table 3a, A2c). But when we look at the predictions (Fig. 25d), curves do not resemble the "drop back" effect in the data. Figure 25 shows predictions from the best algorithms in Table 3a. Some indeed look similar to the plots in Figure 24 but none can fully predict the experimental data. For example, algorithm 9 would also predict one curve with a slope near m=-1 that is seen in the *wide sample1* data (Fig. 24b) but not in the *dense* data set (Fig. 24a). Predictions from algorithms 8a, 10, and 5, although generally reminding of the wide sample2 data in Fig. 24c do not return to low salience at high reference target luminance.

Conclusions from Experiment 9:

Matches of DARK minimum vs. DARK maximum targets strongly suffered from item salience effects in wide blob arrangements. – Salience variations received only partial and singular fits (e.g., algorithm 9 for matches in dense blob arrangements). – Equal-salient luminance variations of DARK minimum targets and DARK maximum targets are about constantly related at a ratio of 1: 0.6.

The test series with *BRIGHT targets* (Experiment 10) were mainly performed on dense blob configurations; only one subject had matched the targets also in wide blob patterns (Fig. 26). Like with the DARK minimum and maximum target matches in Figure 22, the salience variations in reference and test patterns are reversed; *decreasing* target luminance in the reference pattern and *increasing* target luminance in the test pattern would both increase the salience of the according target. But note the differences in linearity. The linear decrease of target luminance in the reference patterns (Fig. 26a, gray lines) is matched by an almost exponential increase of target luminance in the test patterns (Fig. 26b, red and blue data

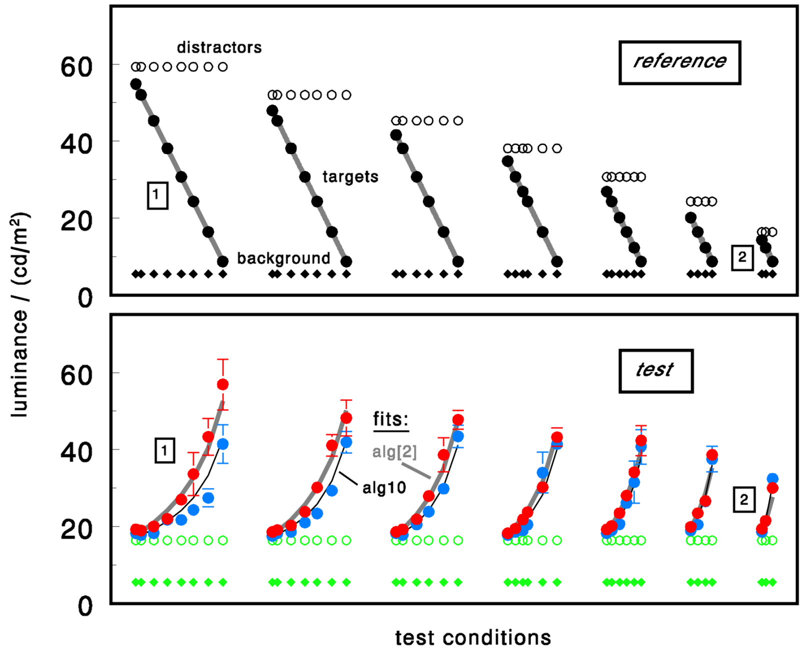


Figure 26. *Equal-salience matches of BRIGHT minimum and maximum targets (Experiment 10).* Symbols and presentation as in previous figures (e.g., Fig.17); reference pattern settings in black, matched test pattern settings in color. Two predictions are shown by gray and thin black lines, respectively (Table 2). Note that to match the linear luminance variations of the reference target, test target luminance had to be increased nonlinearly. Numbers are referred to in the text. For the nomenclature of alg[2] see footnote 3.

points). In contrast, the salience variations with DARK (minimum and maximum) targets were both about linear (Fig. 22). Another difference between BRIGHT and DARK targets is that the BRIGHT target matches with different blob densities were rather similar (Fig. 26b), whereas there have been strong differences, as just discussed, between dense and wide blob densities in the DARK target matches (Fig. 22). Apparently, low item salience did not affect the matches of BRIGHT targets as strongly as it did affect those of DARK targets. The reason is simply the different salience of background-near BRIGHT and DARK targets (see also Fig. 13). BRIGHT targets were presented on dark background, and DARK targets on bright background. Since the target detection threshold varies with background luminance (Nothdurft, 2015; cf. Stevens, 1961), small luminance differences might be sufficient to make a (dim) BRIGHT target appear distinct from the dark background, but larger luminance differences are needed to make a (bright) DARK target appear similarly distinct on the bright background. As a matter of fact, the reference targets at the three lowest luminance settings in Figure 26a (label 2) were all more salient than the reference targets at the corresponding three highest luminance settings in Figure 22. To see and measure item salience effects in the present task, the BRIGHT reference targets should have even more closely approached background luminance.

For MSD analysis, the data from dense and wide blob configurations were averaged and curves fitted by various algorithms. The best result (smallest MSD values) was obtained with algorithm [2][3] on the power of luminance, the predictions of which are plotted into the figure (gray). (Details of these predictions will be given in section C.)

The two-dimensional scatter plots of salience-matched reference and test targets (Fig. 27) show, for the dense blob spacing (Fig. 27c), several similar curves that are shifted in *x* due to the different distractor settings in the reference patterns (dashed lines). Different to Figure 24, discrimination salience now increases from the right to the left, that is, from high luminance values of the reference target (low discrimination salience) to low luminance values (maximal discrimination salience). The according values of the matched test targets indicate a corresponding increase of salience from low to high luminance values. Interestingly, all curves remain separated, which is not quite what we should expect from the experiments in section A. There, the *salmin* algorithm has made the best predictions of both BRIGHT maximum and BRIGHT minimum target matches (Table 2, E5 and G5). In this algorithm, luminance variations are normalized to the luminance span of background and distractors. In

3 For the brevity of presentation, brackets around an algorithm number indicate that this algorithm was performed on the power of luminance (exponent $x=0.33$). The convention is introduced in section C.

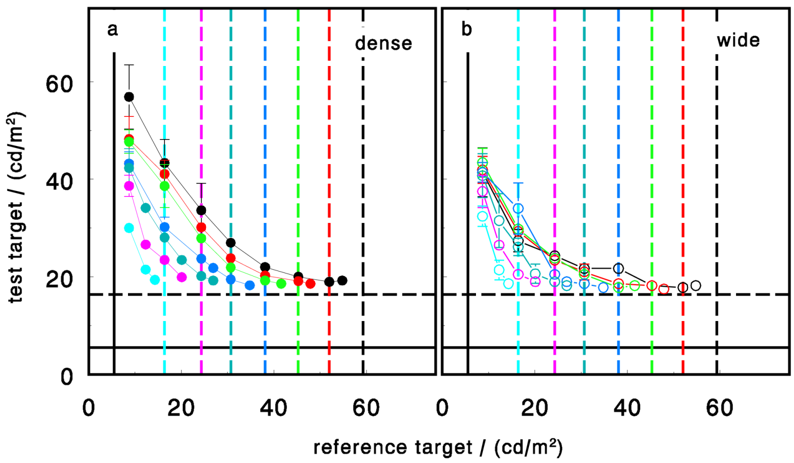


Figure 27. *Target-to-target variations in Experiment 10,* for matches from **a.** dense and **b.** wide blob configurations. Note that curves remain separated in *(a)* but tend to collapse in the left-hand data points of *(b)*. See text for discussion.

Experiment 10 that span was constant in all test patterns, and the variations in reference patterns should have been lost due to the normalization. Thus, if the matches were indeed based on the *salmin* algorithm, the different data curves in Figure 26b should all look similar and the curves in Figure 27 should all intersect in one point (same test target luminance at low reference target luminance; cf. Fig. 28d). This is certainly not the case with the data from dense blob patterns (Fig. 27a). In these patterns, apparently, a large difference of a BRIGHT minimum target to its little brighter distractors (Fig. 26, labels 2) did not make the target as salient as the proportionally identical contrast of the same target to much brighter distractors (Fig. 26, labels 1). This was different in wide blob configurations, for which the different matching curves in Figure 26 look much more similar (blue symbols) and data points of maximal target salience tend to collapse (Fig. 27b, data points on the left-hand side).

Predictions were again made along two lines of computations, using *mixed combinations* of algorithms that made good predictions of earlier matches of the two targets (Table 2, E5 and G5), and *unique algorithms* that were applied to both target types. Only few algorithms provided reasonable predictions of the data (Table 3, C1-D2); these curves are shown in Figure 28 (for more details, see section C).

Conclusions from Experiment 10:
BRIGHT minimum vs. maximum target matches received several good fits including a particularly good one from algorithm [2] (constant ratio on the power|0.33 of luminance)

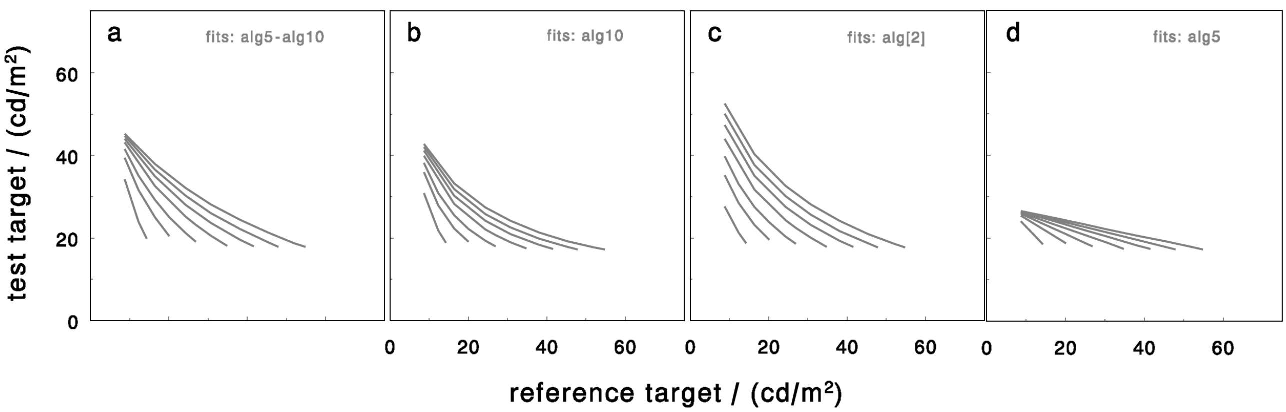


Figure 28. *Various predictions of target-to-target variations in Experiment 10.* Curves show the expected variations in Fig.27 if salience were computed by the indicated algorithms from Table 1; **a.** “mixed combination” of algorithms 5 and 10; **b.** algorithm 10; **c.** algorithm [2] applied to the power of luminance; **d.** algorithm 5 (for details, see text). These algorithms had produced the closest predictions of the experimental data (cf. Table 3, C1-D2).

Experiment 11:
Matches of maximum targets with different contrast polarity *(test series blocks W and WB)*

> *Hints for reading:* The following two experiments are important as they demonstrate the role of power-transforms (Steven's brightness law) in certain matches.

In Experiment 11, BRIGHT targets were matched to DARK targets, and *vice versa;* both targets were presented in maximum target configurations.

Stimuli

All patterns in test series block *W* had similar distractor luminance for reference and test patterns but different backgrounds (Fig. 29). Background luminance was high to let distractors appear dark, or low to let the same distractors appear bright. The salience of DARK targets in the reference patterns had to be matched with that of BRIGHT targets in the test patterns. Block *W* included seven test series that were presented in four runs each with 9 test conditions. All matches were performed on the dense blob raster. At a later stage of the project two additional test series were designed; one covered a slightly reduced collection of previous test series block *W* (series *WB1),* the other contained a set of stimulus conditions in which these matches were reversed, i.e. BRIGHT reference targets were matched by DARK test targets (series *WB2)*. These two new test series included 30 and 26 test conditions, respectively. In every run, all stimulus conditions were presented twice, in pseudo random order, with reference and test patterns presented on both sides of the screen. Every test series was run by three subjects. Test series block *W* was part of the main study and run by the three subjects that had also performed most other experiments of this paper. For the additional test series *WB1* and *WB2*, two of these subjects were replaced by new observers.

Results and Discussion

The stimulus conditions of test series block W were the simplest possible combinations of DARK and BRIGHT targets in maximum conditions, since distractors were identical in the two patterns. Backgrounds had to be different, because they defined the luminance polarity and apparent lightness of items.

Figure 29 shows the matching results for a number of test series in block *W*, together with two straight-forward predictions (gray). With homogeneous blob arrays without distractors it was seen (Nothdurft, 2015) that dark and bright blobs at low and medium contrast appear equal-salient, when their luminance difference to background is the same. If backgrounds differ, however, salience is better related to differences in the power of luminance. If that would similarly hold for the salience matches of targets among identical distractors in Experiment 11, the adjusted luminance settings in Figure 29 should have fallen upon the gray continuous lines indicating predictions from the constant-addition and constant-ratio rules (here identical). This is clearly not the case. If, on the other hand, equal-salient matches were obtained for an equal Michelson contrast of targets to distractors, as reported in other studies (Dannemiller & Stephens, 2001), the data should have fallen upon the dashed gray lines indicating predictions from the Michelson contrast (algorithm 3). While the fit of these latter curves to the data was better, quite a few data points deviated from both predictions, irrespective of whether distractor luminance was high or low and whether distractor-to-background contrast was similar or largely different in the two patterns.

An explanation arises from the scatter plots in Figure 30, in which the salience-matched target settings are plotted against each other (Fig. 30a, filled symbols; matches at and outside the monitor limits are marked with gray fillings). The various data curves do not run strictly perpendicular to the line of identity (dotted) but are slightly bent upwards indicating that the luminance increase with the BRIGHT test targets (plotted upwards) was larger than the according luminance decrease of the equal-salient DARK reference targets (plotted from right to left), as already seen in Figure 29. The general course of data curves becomes more obvious when the test conditions are exchanged. In principle, equal-salience matches should be exchangeable, that is salience matches of test patterns to reference patterns should be similar to salience matches of reference patterns to test patterns.

Reversed matching of test series block *W* was not included in the original test series and could not later be added, since two subjects were not anymore available at this stage of analysis. In Figure 30a, instead, the data were mirrored (open symbols), as if reference and test settings

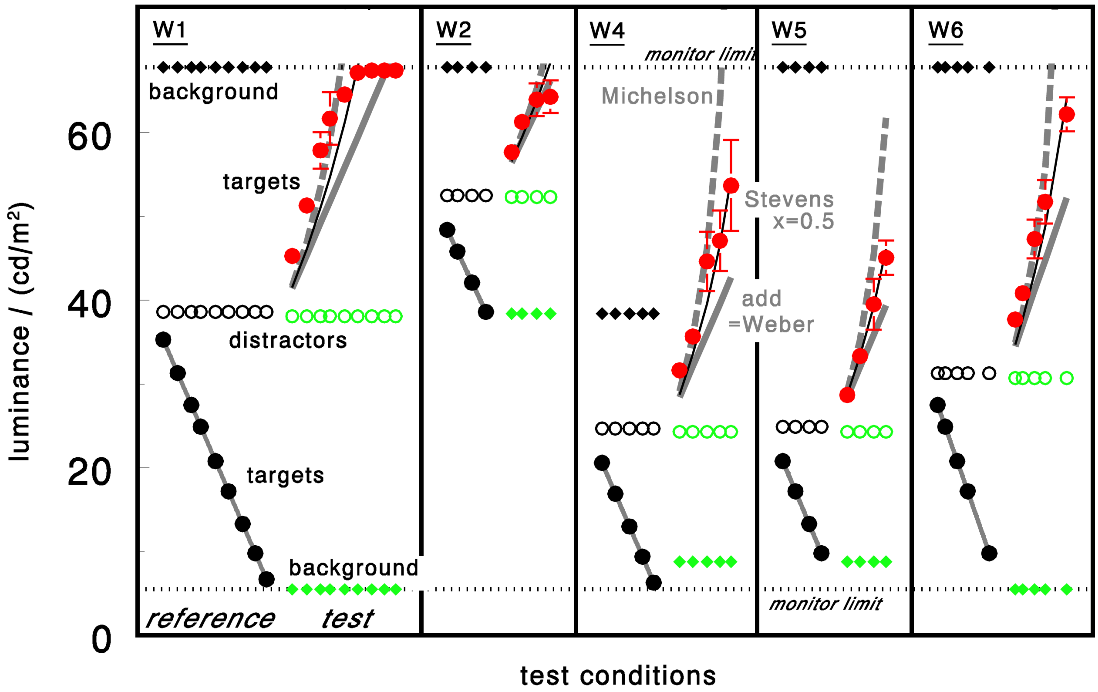


Figure 29. *Examples of equal-salience matches of BRIGHT to DARK maximum targets in Experiment 11 (test series W);* symbols and presentation as in Fig.3. Reference pattern settings (black) and matched test pattern settings (color) are plotted side by side; only dense blob configurations were tested (red). Distractors in the two patterns had the same luminance (aside from minor stray-light effects from the different backgrounds) but looked differently, dark or bright. Three predictions are shown with the data; constant addition ("add"; continuous gray lines) equivalent to equal Weber contrast in these tests; equal Michelson contrast (dashed gray lines); and equal differences in the power of luminance, exponent $x=0.5$ ("Stevens"; thin black lines). Matches are not well predicted from Weber or Michelson contrast. Dotted black lines indicate monitor limitations.

were exchanged. In Figure 30b, however, data from similar matches of the original test conditions (now test series *WB1*) and of new test conditions in reversed matches (test series *WB2)* are shown; these tests were performed by new subjects (and the author) at a later stage of the project. Both graphs in Figure 30 suggest a bent course of data curves as it would be obtained when parameter variations are not linear but follow a power function. For luminance, exponents of sensory power functions are reported to lie between 1.2 (lightness in reflections from gray paper) and 0.33 (brightness of a 5° target; Stevens, 1961); the according variations of equal target-to-distractor differences are schematically shown in Figure 31. Brightness sensation for a point source target (perhaps more comparable to the targets in the present study) follows a power function with the exponent $x=0.5$ (Stevens, 1961). An exponent of $x=1.0$ represents the linear case.

If sensitivity would be linear and identical increments and decrements of targets from the distractors would be seen as equally salient, the scatter plots in Figure 30 should reveal straight lines at slopes $m=-1$ (Fig. 31a). But if sensitivity would follow a power function of luminance, equal sensations of increments and decrements would correspond to different luminance intervals and data curves of equal-salient targets should be bent similar to those in Figure 31b and c. A rather good fit of the data in Figure 29 is achieved with the power transform (exponent $x=0.5$) of luminance.

How would these curves change when distractors in reference and test patterns are not identical? Additional test series (not reported here) that included test conditions with *different* distractor settings revealed reasonable fits for constant-addition on power-transformed luminance data (as schematically predicted in Fig. 31d). In extreme combinations, however, such as low distractor contrast in

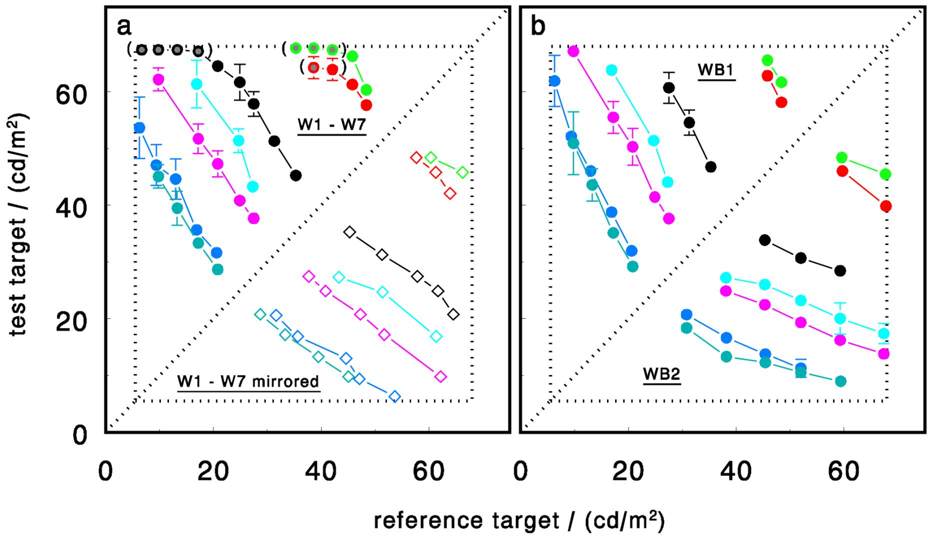


Figure 30. *Target-to-target variations in Experiment 11 (cross-polarity matches).* **a.** Test series block *W.* Matches of BRIGHT test targets to DARK reference targets (filled circles; upper left) among identical distractors; the reversed conditions (BRIGHT reference targets, DARK test targets) are mirrored from these data (open squares; lower right). **b.** Test series block *WB.* In a later repetition of test series *W* both matching directions were tested; DARK reference and BRIGHT test targets (upper left), and BRIGHT reference with DARK test targets (lower right). Curves suggest a continuous transition of curves, in which the luminance difference of BRIGHT targets to distractors increases faster than that of equal-salient DARK targets. Colors distinguish different test series; matches were performed only on dense blob patterns. Dotted lines indicate identity (i.e., zero increments and decrements from distractors) and monitor limits, respectively. Gray data points in *(a)* are invalid because of these limits and are not mirrored. Data points within brackets were not included in test series *WB1* in *(b)*.

one pattern and high distractor contrast in the other, matches could deviate from these predictions, and no final conclusions can yet be made. A preliminary analysis of these deviations, however, suggests that they might be also affected by brightness induction effects (cf. Blakeslee & McCourt, 2013; Kingdom, 2011).

Thus, despite the different findings for equal-salience matches with other stimulus combinations, equal-salience matches of BRIGHT and DARK maximum targets are closely predicted by constant differences of power-transformed target and distractor luminance settings. According to Stevens (1961), an exponent of $x=0.5$ would be adequate to describe brightness perception, and good matches were indeed found for power functions with this exponent (cf. black line predictions in Fig. 29).

Conclusions from Experiment 11:

Cross-polarity matches of DARK and BRIGHT maximum targets among identical distractors (dark or bright) were best predicted by constant differences on a power function of luminance (exponents $x=0.5$ and $x=0.33$).

Experiment 12:
Matches of minimum and maximum targets across different luminance polarities *(test series block WM)*

The observations of Experiment 11 suggested a new block of test series that were only partly included in the original test collection. These test series were supposed to combine reference targets in *minimum* configuration with different categories of test targets, when reference and test distractors were identical. Of the possible combinations only matches of (similar) minimum and maximum targets had been tested in Experiments 9 and 10.

Stimuli

The four possible combinations of minimum and maximum targets (Fig. 32) were tested in four experimental runs, each containing two test series with 5 and 7 test conditions, respectively. In all conditions, distractors in the two patterns were identical. All matches were performed with dense and wide blob configurations.

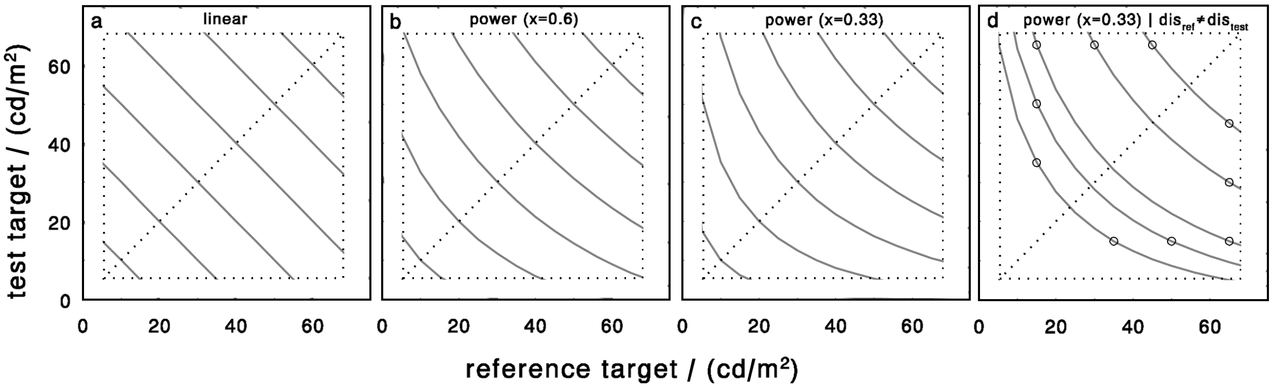


Figure 31. *Schematic drawing of target variations in Experiment 11.* **a.** Equal increments and decrements when luminance scaling is linear. **b-c.** Nonlinear variations when increments and decrements are equal in the power of luminance (different exponents). Curves in Fig.30 are better described by *(b)* or *(c)* than by *(a)*. **d.** For dissimilar distractor settings in reference and test patterns (open circles) curves are shifted but remain their principle form.

Experimental procedures were identical to those in the previous experiments; test targets had to be adjusted to appear equal-salient to the reference target. Matches were performed by three subjects, two of which had not contributed to the main pool of experiments of this study.

Results and Discussion

The simple stimulus schemes in Figure 32 give an overview of how these various matches may be looked at. When DARK targets in minimum configuration are combined with DARK targets in maximum configuration *(WM1)*, or BRIGHT targets in minimum with BRIGHT targets in maximum configuration *(WM4),* both patterns may have the same background, and target salience would increase with luminance changes in opposite directions. Both combinations were already tested in Experiments 9 and 10 but with different distractors.

The remaining two combinations, DARK targets in minimum configuration combined with BRIGHT targets in maximum configuration *(WM2)*, or BRIGHT targets in minimum configuration combined with DARK targets in maximum configuration *(WM3)* were not yet tested but the identical distractors will create a situation in which targets are identical and have to be matched among identical distractors. This is, in principle, an identity match which might only be “disturbed” by the different backgrounds on which distractors and target are presented. It would be interesting to see if matches indeed reflect these identities or are strongly affected by the different backgrounds and hence the different lightness of items.

Matches are shown as scatter plots in Figure 33. In test series WM1 and WM4 (Fig. 33a), data points indeed follow the predictions of constant addition on power-transformed luminance scales (gray curves) and only deviate from these curves for matches close to the

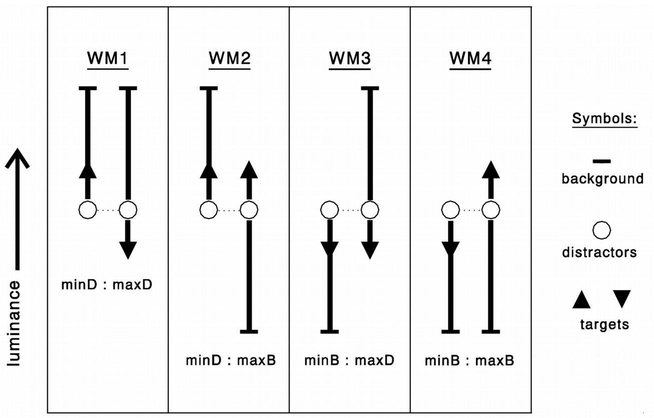


Figure 32. *Schematic drawing of target conditions.* Background luminance is indicated by short horizontal lines, distractor luminance by open circles, and target luminance by arrows (pointing into the direction of increasing discrimination salience). The scheme reveals principle similarities in the various combinations of DARK (“D”) and BRIGHT targets (“B”) in maximum (“max”) and minimum (“min”) configuration. These combinations are tested in four test series (WM1-WM4) in Experiment 12.

ISSN:2364-3641

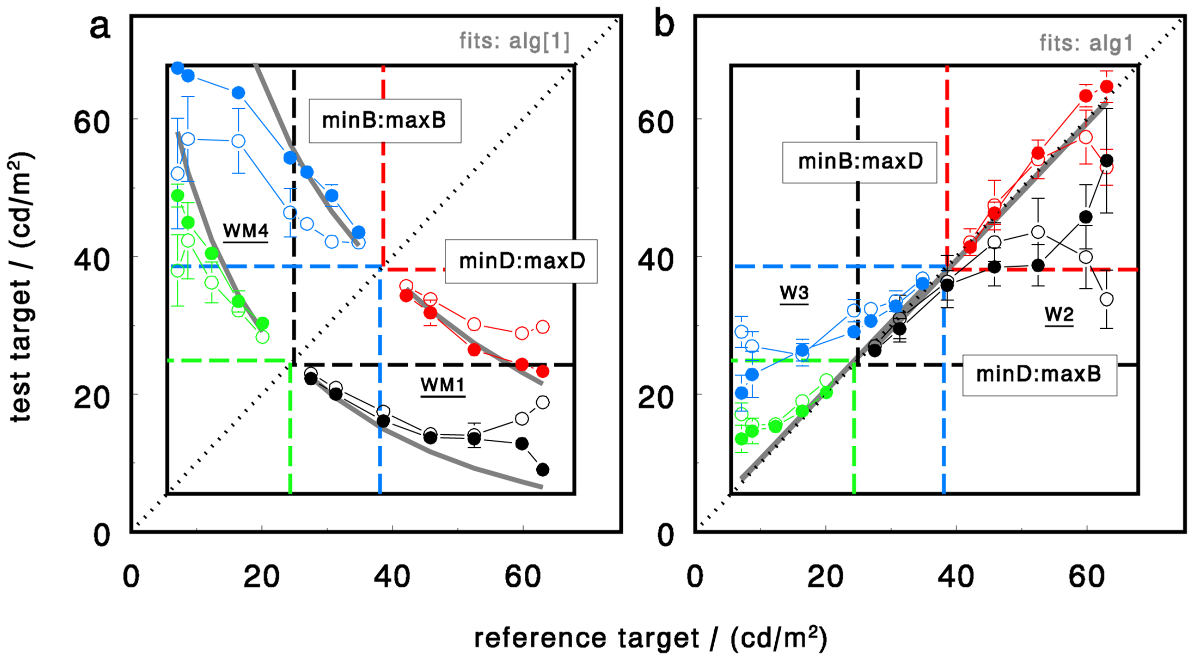


Figure 33. *Scatter plots of target-to-target variations in Experiment 12.* **a.** Target variations in opposite direction (increments versus decrements); **b.** identical target variations. Curves plot equal-salience matches of the target combinations sketched in Fig.32. For each combination, two test series were performed (different colors) on dense (filled circles) and wide blob patterns (open circles); the according distractor levels are shown as dashed lines in the same colors. Backgrounds were set to 5.5 and 68 cd/m^2, respectively (continuous black lines). For target variations in opposite directions (a), equal-salience matches tend to follow a power function with an exponent near $x=0.33$ (fitted gray curves). For similar target variations *(b)*, data indeed reveal identity matches (fitted gray curves), despite the different lightness of targets. Only when target-distractor differences are increased, most curves deviate from the initial fits.

luminance limits of the monitor which had served as background settings of the various targets (continuous black lines) or at high target-to-distractor contrast (data points far away from the mid line), where discrimination salience might have interfered with low item salience in the wide blob patterns (open symbols). As expected, the curves look by and large similar to the curves obtained for matches of DARK and BRIGHT maximum targets in Figure 30. Quite differently, but also as expected, most matches in test series WM2 and WM3 (Fig. 33b) fell close to or exactly upon the identity line (gray line), that is, they were matched identical, despite the different backgrounds and despite the different target types that were here combined. Again, there were deviations, particularly for matches in wide blob configurations (open symbols), when DARK minimum targets approached the bright background and item salience vanished (data points on the right-hand side). Altogether, thus, the findings from Experiment 12 confirm and generalize the observations from the previous experiments.

There are additional deviations from the expectations in Figure 33, even for matches in dense blob configurations. These deviations are likely caused by the different background settings in the two patterns. Note that all curves start close to the expected prediction lines for small target-to-distractor contrast (data points near the intersection of the according distractor settings marked by dashed lines in the same color). While some data curves remain near these lines with increasing target-to-distractor contrast, others deviate from them. This is particularly obvious for the black and blue data point series. In the previous experiments, only DARK maximum target matches were (largely) independent from background, whereas all BRIGHT target matches were scaled (normalized) to the according background-distractor luminance span.

Conclusions from Experiment 12:
For small to medium luminance variations, equal-salient minimum and maximum target matches among identical distractors follow the same general rules, independent of the apparent brightness and lightness of the targets. Targets that vary in opposite directions (increments and decrements) follow the power function of luminance. Targets that vary in the same direction (both increments or both decrements) represent identity matches.

Experiment 13:
Confirmation of the maximum-minimum salience difference *(test series block WZ)*

Experiment 12 (and also Experiments 9 and 10) included the maximum-minimum paradigm described in the Introduction (Fig. 1). The darkest or brightest item among less dark or less bright distractors (the *maximum* target) is more salient than a single distractor among many of these targets (the *minimum* target condition; cf. Nothdurft, 2006a). Thus, a simple exchange of the luminance settings of targets and distractors (which does, of course, not change their luminance difference) should affect the salience of the target. To directly measure the salience difference in the testing paradigms of the present study, test distractors in Experiment 13 were given the luminance settings of reference targets, on identical backgrounds, and test targets were then adjusted for equal salience. If there were no difference and salience were simply based on the (identical) luminance contrast of targets and distractors, the salience-matched test targets should be adjusted to the same luminance settings as the reference distractors so that luminance settings of targets and distractors were simply exchanged.

Stimuli

This last experiment included two test series, one for DARK and one for BRIGHT targets, each with eight stimulus combinations. Three subjects served as observers; two of them had not been involved in the main data collection of the study.

Results and Discussion

Figure 34 shows the luminance settings of reference and matched test patterns; reference patterns represent the minimum target configuration. In both test series, reference distractors (open black circles) were held constant and reference targets (filled black circles) were systematically varied between background and distractor settings. (In the figures, these conditions are shifted so that reference target settings fall upon a straight line.) Test patterns were presented on the same backgrounds, with distractor and target luminance settings virtually exchanged. That is, test distractors (green open circles) were set to the luminance level of reference targets (black circles), and test targets (to be adjusted) were initially set to a luminance just below (DARK targets) or above that of test distractors (BRIGHT targets). Blue circles represent the matches obtained with the wide blob raster, red symbols those with the dense blob raster. If salience computations were symmetrical, all test target data should have fallen upon the symbols of reference distractors, which obviously was not the case. Test targets were generally matched to smaller target-to-distractor differences than those in the according reference patterns. Thus, to make maximum and minimum targets equally salient, the target-to-distractor contrast of the maximum targets had to be decreased. In turn, if the contrast were not decreased but set to the (larger) luminance difference in the reference pattern, maximum targets had been more salient than the minimum targets, which is, in fact, the description of the maximum-minimum paradigm.

It is interesting to look at how the test target settings deviate from the "symmetrical" case, in which target and distractor luminance settings were simply exchanged. With increasing target-to-distractor contrast in the reference patterns (i.e., moving from the right to the left in the graphs of in Fig. 34), test targets first follow virtual straight lines at a smaller target-to-distractor luminance difference than that in the accompanying reference pattern. Thus, with increasing target-to-distractor contrast, the deviations from the symmetrical case increase. For DARK targets, the matches in wide (blue symbols) and dense blob patterns (red symbols) were initially similar; for BRIGHT targets, however, these matches differed. When the luminance difference of target and distractors in the reference patterns is further increased (data points towards the left), reference targets will finally approach background luminance and *item salience* in wide blob configurations will decrease. Here, different response variations are seen with the DARK and the BRIGHT targets. With DARK targets in wide configurations, the matched test target-to-distractor contrast was further decreased indicating a strong and over-proportional reduction of target salience; in dense configurations, the matched target-to-distractor contrast remained about constant and only the last matches with very bright reference targets (near background level) were a little uncertain (cf. the larger s.e.m. in these conditions). With BRIGHT targets, however, item salience of reference targets near background luminance was not so strongly reduced and salience-matches continue to follow the initial course of reduced contrast (blue line). Neither in wide nor in dense configurations did the test target matches systematically approach the test distractor settings, as it was the case with DARK targets. We have seen a similar difference between DARK and BRIGHT minimum targets in Experiments 9 and 10 (Figs. 23 and 26). Due to the different background luminance in these patterns and the dependence of detection threshold on luminance, BRIGHT reference targets close to (the dark) background suffer less from diminished item salience than DARK reference targets close to (the bright) background. BRIGHT reference targets had to be even closer to background luminance to evoke a similar item salience reduction as seen with DARK targets.

The data are re-plotted in the scatter plot of Figure 35. Note that the sequence of test conditions with DARK targets is then reversed (test series *WZ1)*. Reference targets and test distractors are located on the oblique mid lines (dotted lines); for strictly symmetrical matches, test targets

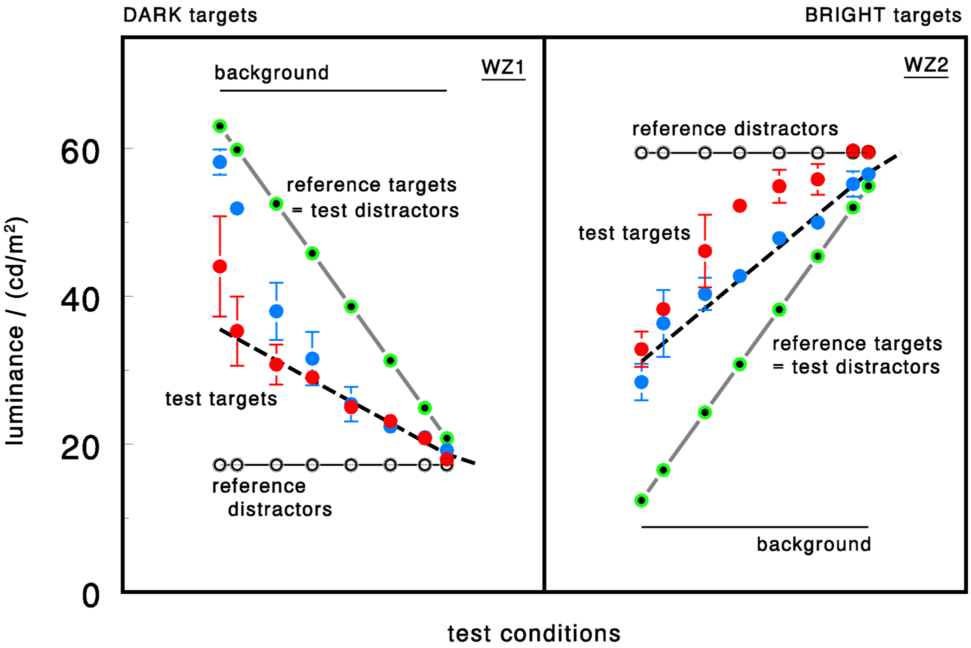


Figure 34. *Test of the maximum-minimum paradigm (Experiment 13).* In two test series, one with DARK and one with BRIGHT targets, the different salience of targets in maximum or minimum configuration was measured. In the original paradigm, target and distractor luminance settings are simply exchanged (on identical backgrounds); the target in the maximum configuration appears then more salient than the target in the minimum configuration (cf. Fig.1). In experiment, test distractor (green circles) and reference target levels (filled black circles) were set identical and test targets in the maximum configuration (blue and red circles) were adjusted to equal the salience of the reference targets in the minimum configuration. Backgrounds (horizontal black lines) were constant and identical. Test conditions are shifted in x so that reference target and test distractor variations line up. Test targets were always adjusted to smaller target-to-distractor differences than reference targets, indicating that the latter were less salient than a maximum target with exchanged luminance settings and the same target-to-distractor luminance difference in the original paradigm condition. Dashed black lines emphasize initial deviations from hypothetical symmetry.

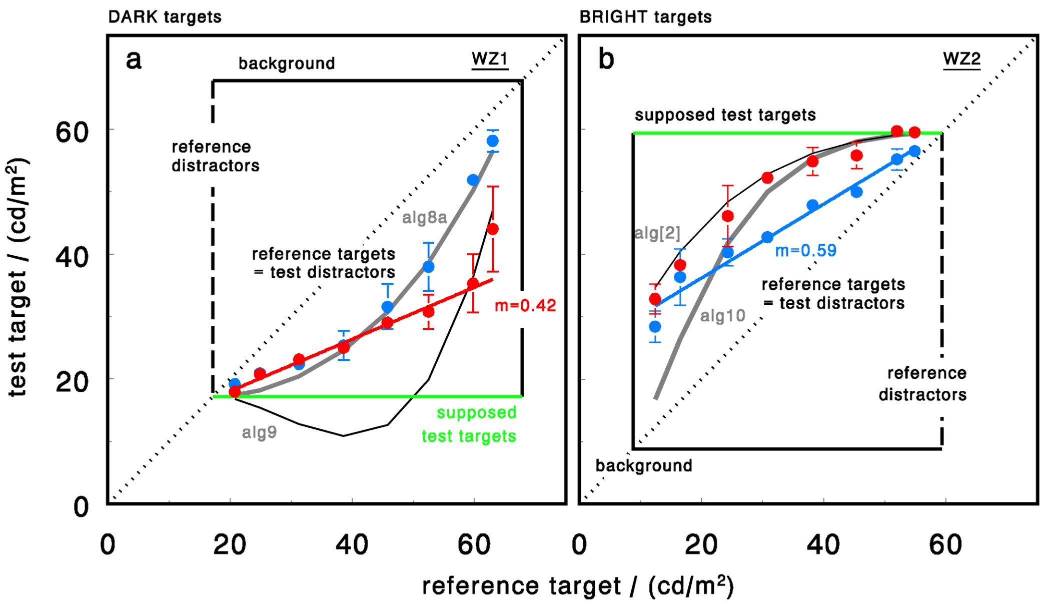


Figure 35. *Target-to-target variations in Experiment 13.* **a.** DARK targets; **b.** BRIGHT targets. Data re-plotted from Fig.34; constant settings are indicated by vertical and horizontal lines. (Note that the sequence of test conditions in series *WZ1* is here reversed compared to Fig.34). Reference targets and test distractors lie on the oblique midline (dotted); in the case of perfect symmetry, test targets should have fallen upon the green lines. Deviations (red and blue data points) show the maximum-minimum paradigm. Red and blue lines are regression lines through the data (the rightmost red data point in *WZ1* was not included); their slopes are given aside. Thick gray and thin black curves show different predictions of the data (see text).

should have fallen upon the green lines. Red and blue lines are regression lines through the data (in series *WZ1* without the right-most red data point). It is interesting to compare the red line obtained for DARK targets (*m=0.42)* with the parallel lines obtained in the related task of Experiment 9 (Fig. 24a). In both experiments, DARK minimum and maximum targets were matched for salience, but test conditions were slightly different. In both test series (Fig. 24a and Fig. 35) reference target luminance was systematically varied between the constant reference distractor settings (dashed vertical lines on the left) and the constant background (continuous vertical lines on the right). But while test distractors were constant in Figure 24a (horizontal dashed line), they varied between conditions and fell upon the dotted line in Figure 35. If Figure 35 were rotated so that the test distractor settings become similar to those in Figure 24a, the tilted dotted line *(m=1.0)* would turn horizontal *(m=0.0)* and the fitted red line *(m=0.42)* into a line similar to the lines in Figure 24a *(m=0.42–1.0=-0.58)*. Both experiments thus reveal a similar ratio of equal-salient target variations of DARK minimum and maximum targets. The ratio is not 1 but about 0.6, that is, the luminance variation of a DARK minimum target would correspond in salience to only about 3/5 of that luminance variation in the matched DARK maximum target.

A similar comparison could be made for BRIGHT minimum and maximum targets. But note that the corresponding target variations in Figure 27 are not linearly related (see also Fig. 26) and straight regression lines would only fit subsections of each curve, as is also the case with the regression lines in Figure 35 (BRIGHT targets, red and blue lines).

How would the best predictions from Experiments 9 and 10 fit the new data from Experiment 13? The unique algorithms that had made best predictions in the earlier experiments were algorithms 8a and 9, for DARK targets in *wide* (sample2 data) and *dense* blob configurations, respectively (Table 3a, A2a and A2c), and algorithm 10 and [2] for BRIGHT targets (Table 3, C2 and D2). These predictions are superimposed in Figure 35 (gray and thin black lines). Two of them fit the data quite well, but the other two made partly wrong predictions. According to the predictions from algorithm 9, for example, DARK targets at intermediate target-to-distractor contrast had to be adjusted even below the reference distractor luminance, indicating that maximum targets at these conditions were *less* salient than minimum targets with exchanged luminance settings, which clearly was not the case. Algorithm 10 suggests that the loss of item salience for BRIGHT targets near the background should be much stronger than found.

Conclusions from Experiment 13:

The *maximum-minimum paradigm* was confirmed; minimum targets are less salient than maximum targets with exchanged luminance settings. Equal-salient luminance variations of DARK minimum targets corresponded to about 60% luminance variations in equal-salient DARK maximum targets.

Discussion of section B

Hints for reading: The Discussion summarizes the major findings of section B: (1) the advantages of using a constant "salience meter" for target comparisons; (2) the occurrence of different salience aspects such as discrimination salience and item salience; and (3) the important role of power functions in cross-polarity target matches. Discussion ends with an illustration of the maximum-minimum paradigm in different blob densities.

While the experiments of section A had identified rules of how targets must be changed to remain their salience when background or distractors change, the experiments of section B help to relate the salience of *different* targets to another. Some observations were confirmatory and underlined the validity of earlier conclusions; others are new and provide information that could not be concluded from the previous section.

Perhaps the most important advantage of these new experiments was the fact that they provided an in-dependent *measure of salience*. By using fixed luminance settings for the test pattern, not only the salience of different targets could be directly compared but also the measures of different reference patterns, even from different test runs, could all be related to the same measure and thus be virtually compared. Another important finding came from target *matches across luminance polarities*, i.e. from DARK and BRIGHT targets among identical (but differently looking) distractors. Are these matches solely based on the luminance settings of targets and distractors or do they also depend on the perceived lightness and brightness of items (cf. Maertens & Wichmann, 2013; Moore & Brown, 2001)? While certain matches, like those in Experiment 12, seemed to reveal little influence from perceptual categories but reflected the identical luminance settings, at least for small differences of targets and distractors (cf. Fig. 33), the luminance settings of the backgrounds (which had caused the percept of dark or bright stimuli) turned out to be essential in other matches. For example, if the backgrounds in Experiment 11 had not been different but had both been either dark or bright, the resulting matches of minimum and maximum targets would have followed different rules (cf. Exp. 9-10) than the performed matches of DARK and BRIGHT maximum targets. A third major observation from the experiments in section B is the important *role of power-transforms* in matches of targets at different luminance polarity. Matches of BRIGHT and DARK targets could largely be predicted when luminance scales were power-transformed with an exponent $x=0.5$, whereas most predictions on a linear scale could not explain the data. The role of power-transforms in salience computation was less obvious in the previous matching experiments of this study. And last not least, salience variations in DARK minimum targets could be related to those in DARK maximum targets (Exp. 9 and 13). Matches revealed a *systematic reduction to about 60%*, that is, luminance variations of the minimum target would be equal-salient to smaller luminance variations of the maximum target.

In addition to these new observations, the experiments of section B have confirmed major observations from section A. Matches of DARK or BRIGHT maximum targets with the new salience measure (Exp. 7 and 8) revealed the same salience "rules" as Experiments 1 and 4. And the observation that target salience is based on a combination of discrimination and item salience, which may change independently in certain stimuli, was strongly confirmed in Experiment 9. When the salience of a DARK target in minimum configuration was compared with a fixed salience meter (i.e., with constant test pattern settings), salience did not continuously increase with increasing target-to-distractor differences but rapidly diminished again when target luminance approached background luminance. It was observed (and explained) that item salience effects should not occur with the dense blob raster and should be less pronounced with BRIGHT targets on a darker background.

Different aspects of salience: item and discrimination salience

The observation itself is not surprising; a target that is barely visible cannot be very salient. But it underlines the different aspects of target salience that must be distinguished. *Item salience* makes a target stand out from background and *discrimination salience* makes a target stand out from other items. In maximum target configurations the item salience of a target is always larger than the item salience of distractors, and it is therefore mainly the discrimination salience that makes the target more conspicuous than other objects in the scene. The item salience of targets in such configurations can only be low when the item salience of distractors is even lower.

For targets in minimum configurations, however, or targets among distractors of opposite luminance polarity, the two salience components may vary independently, and a target with high discrimination salience that would be potentially highly conspicuous among the distractors, could nevertheless display very low item salience and hence might not be conspicuous at all. To my knowledge such a distinction of different salience aspects, although important, has not yet systematically been made in the literature. It was reported (but not named) in Nothdurft (2000) where matches of indistinguishable targets (zero discrimination salience) did not measure zero salience. The distinction of different salience components may help to solve hitherto open questions or contradictions in salience research. For example, a target embedded in identical targets is not entirely non-salient, although discrimination salience would be zero. An important aspect of item and discrimination salience is the distance to other objects. If accompanying items are presented closely to the target, the distinction from background vanishes and discrimination salience is becoming the most important aspect of salience. In agreement with this view, item salience was so far mainly noticed in sparse target and distractor arrangements (cf. Nothdurft, 2000).

Salience matches across luminance polarities

The combination of DARK and BRIGHT targets (Exp. 11 and 12) cannot be obtained from illumination changes on purely reflective surfaces, and thus there is no "natural" algorithm that could be expected to predict equal-salience matches of these targets. Another starting point for equal salience computations of DARK and BRIGHT targets would be Weber's law that the just discriminable luminance difference depends on background luminance ($\Delta I/I$ = constant), no matter whether the difference is added or subtracted. For supra-threshold differences, however, this rule is often violated. Dannemiller and Stephens (2001), for example, found that children rate BRIGHT targets as less conspicuous than DARK targets with the same Weber contrast ($\Delta I/I$) on identical backgrounds, and concluded that the Michelson contrast ($\Delta I/(I+\Delta I)$) might be better suited to predict equal salience[4]. In the present study (Exp. 11), matching performance was better related to the power of luminance and the Michelson contrast also produced only partly good predictions of some matches (cf. Fig. 29; Table 4). However, when item salience rather than discrimination salience is matched (as in the accompanying study on uniform blob arrays; Nothdurft, 2015), equal salience is not predicted by the Michelson contrast, but by the constant-ratio rule (Weber contrast). Dannemiller and Stephens (2001), too, found ratings congruent with the Weber contrast for small luminance differences. Altogether, this leaves us with apparently heterogeneous and multiple salience computation rules. The *item salience* of dark and bright blob arrays follows the constant-ratio rule (Weber contrast) when blobs are presented on the same backgrounds. But when presented on different backgrounds, equal-salience is better described by constant differences on a power-transformed luminance scale (Nothdurft, 2015) as in Stevens' brightness law (Stevens, 1961; Rudd & Popa, 2007)). The *discrimination salience* of DARK and BRIGHT blobs among identical dark and bright distractors, however, clearly follows the power of luminance (Fig. 30) and only partly the Michelson contrast (Dannemiller & Stephens, 2001).

Michelson contrast and the power function of Steven's brightness law (exponent $x=0.33$) show a principally similar course with increasing luminance differences but differed in the fits of Figure 29. According to Table 4, the supra-threshold matching performance in Experiment 11 was better predicted by Steven's brightness law (algorithm [1]) than by the Michelson contrast (algorithm 3). This was also the case for power transforms with the exponent

4 With $\Delta I=(I_{max}-I_{min})/2$ and $I=I_{min}$, their definition of Michelson contrast is equivalent to the definition used here, $(I_{max}-I_{min})/(I_{max}+I_{min})$; see, e.g., Peli, 1990).

x=0.5 (cf. the different fits in Fig. 29). Since bright and dark targets on the same background cannot have equal brightness, it is interesting to note that Steven's brightness law thus also holds for equal-salience matches across contrast polarities.

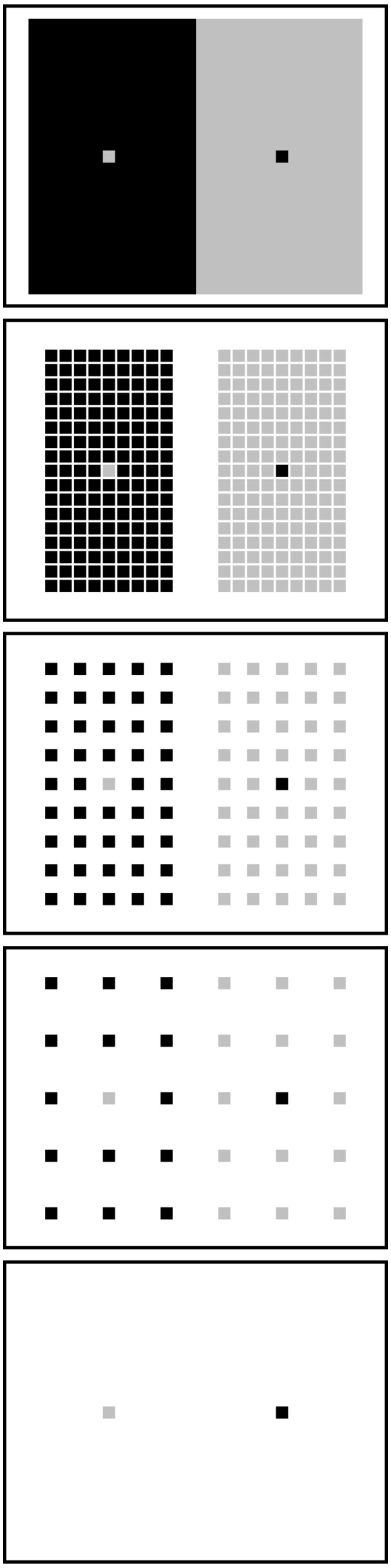
The maximum-minimum paradigm

All matches of minimum and maximum targets (Exp. 9, 10, and 13) confirmed the maximum-minimum paradigm (cf. Fig. 1). In this paradigm a single item with strongest item salience (i.e., a target in maximum configuration) stands out more strongly from less salient distractors than a single less salient distractor (i.e., a target in minimum configuration) stands out from many target items (Nothdurft, 2006a). In the originally proposal, target and distractor luminance settings were simply exchanged between the two conditions. In Experiment 13 designed to *match* the two targets in salience, the contrast of the maximum target had always to be reduced to lower its salience; it never was increased above the item contrast of the reference distractors (cf. Fig. 34). The need to reduce the salience of the maximum target to make the two targets look equal-salient was seen with both DARK and BRIGHT targets and thus confirms the general validity of the proposed maximum-minimum paradigm.

An interesting issue would be the possible influence of blob density in these comparisons. One could expect that in patterns with densely arranged blobs it is not the item salience of targets and distractors that is compared but the discrimination salience (target-to-distractor contrast), which should be based on identical differences for the two targets. Couldn't it be that for a certain blob density discrimination salience would be so dominant that the perceived difference in the maximum-minimum paradigm would disappear? The data in Figure 34 indicate that this will likely not happen. In both curves, obtained with the wide and dense blob raster, salience-matched maximum targets displayed smaller luminance differences than the according maximum targets.

The salience variations with blob density are illustrated in Figure 36. Note that the luminance settings in all patterns are identical and that target and distractor settings in the left- and right-hand patterns are exchanged to illustrate the maximum-minimum paradigm. In one extreme case (uppermost patterns), distractors are so densely packed that the background is not seen. In the other extreme (bottom), items are so widely spaced that all distractors fall outside the pattern and only single targets are shown. In the other patterns, targets and distractors are show in various densities.

How does the relative salience of the two targets change with density? Let's start with the uppermost case, which is somewhat special. In principle, the item salience of a single target on background is related to its Weber contrast (Nothdurft, 2015). Since the increment and decrement of the two targets to their backgrounds are the same, the ratio to backgrounds would be important. According to the Weber contrast, the left-hand BRIGHT target on the dark background should appear more salient than the right-hand DARK target on the bright background. Several studies (e.g., Dannemiller & Stephens, 2001) might predict the opposite, as they have found decrements being rated more salient than equal increments – on the same background. For *different* backgrounds, however, salience is better described by differences in the power of luminance (Nothdurft, 2015, Exp. 7), which should then make both targets equal-salient. That is also the impression most observers have when looking at the top row patterns of Figure 36. When moving down to the other patterns, the impression of equal salience disappears and the maximum targets on the right-hand side become clearly more salient than the minimum targets on the left. This salience difference remains until finally, in the bottom pattern, single items are compared. The backgrounds here are identical and both targets are DARK, thus target salience should correlate with the target-to-background Weber contrast (Nothdurft, 2015), which is, of course, larger for the target on the right-hand side. The figure nicely illustrates the interplay of the two salience aspects, item salience and discrimination salience. Targets in the bottom pattern only show item salience, which is different for the two targets. Targets in the pattern on top only show discrimination salience which is apparently similar for the two targets. (It is, of course, item salience, which lets the targets stand out here, but from construction, the pattern displays the luminance difference to distractors, not to the background.) From top to bottom, discrimination salience diminishes while item salience becomes more dominant. Whether or not this transition is linear, is not quite clear yet. Note that matches also change from comparing a bright target (left) and a dark target (right), in the top rows, to comparing two dark targets, in the lower rows. Some people find the salience of the left-hand target in certain densities particularly low (e.g., 3rd or 4th row in Fig. 36).



Whether that is a true observation or an impression caused by the rapid loss of discrimination salience in comparison with the patterns above needs to be studied. However, the center-surround structure of neurons in early processing stages of the visual system ("lateral inhibition") should allow for local sinks and troughs of (minimum) target salience at certain blob densities.

Conclusions from discussion:

Stevens' brightness law also holds for salience matches across contrast polarities, if distractors are identical in luminance. Under these conditions, small target variations in minimum and maximum targets are matched identical. – The maximum-minimum paradigm holds for all blob densities.

Summary of Section B

Section B presented equal-salience matches of different target types along three major lines of experiments.

1. Targets in maximum and targets in minimum configuration (cf. Fig. 1) were matched against a constant salience meter, which did not only identify the rules of equal salience in these stimuli but also provided a useful basis for measuring salience differences between targets. In certain configurations, the different effects of *item salience* and *discrimination salience* were again seen.

2. Targets at different luminance polarity to background confirmed the importance of Steven's brightness law in these matches. In addition, quite different target types were often seen to be matched according to their luminance settings and not their lightness.

3. The maximum-minimum paradigm was confirmed with items. In all presentations, minimum targets were less salient than the corresponding maximum targets. In a demo it was shown that this difference holds for all raster widths.

Contrary to the experiments in section A, matching performance in section B could sometimes not be well predicted from simple algorithms, which suggested me to study salience computation in a larger variety of model algorithms in section C.

Figure 36. *The maximum-minimum paradigm in different blob densities.* In the packed case (top), both targets are about equal-salient (discrimination salience). With decreasing blob density, discrimination salience diminishes and only item salience remains. In the singles case (bottom), the two targets differ notably.

C: COMPUTATIONAL MODELS

Hints for reading: This section is very computational – and may look more frightening than it should. Various algorithms are tested whether they can predict the experimental data. Experiments from section A are analyzed in part I, experiments from section B in part II. The comparisons of predictions with the data are very detailed; skip them if you feel bored – best predictions will later be summarized. Particularly in part II, there is little generalization – except that power-transforms are (mainly) important for cross-polarity matches.

The failure to find a unique algorithm for equal-salience matches in *all* configurations suggested a more general evaluation of computational models. The idea was to define a large set of possible salience algorithms and test which of them can, or cannot, predict the experimental data. The original hope was to identify *one* algorithm that would predict performance in all various matching tasks.

Selection of proposed salience algorithms

An overview of the tested algorithms is given in Table 1. It includes algorithms that were intuitively considered to represent possibly valid models of salience computation (cf. Nothdurft, 2015), like the constant-addition principle (algorithm 1) and the constant-ratio principle (algorithm 2). These algorithms were already introduced in sections A and B. Other algorithms were added because they appeared to fit the experimental data in some experiments, like the *salmin* algorithm 5 also introduced in section A, or because they had been suggested in the literature on related topics (e.g., transparency), like the *Singh* algorithm 8 (Singh and Anderson, 2006). Quite a few algorithms were included because they represented, in principle, plausible combinations of target, distractor, and background luminance; some of them are only listed for completeness. For example, when adding an algorithm that computes the ratio of target-to-distractor and distractor-to-background Weber contrasts (algorithm 9), it seemed reasonable to include also an algorithm that computes the same ratio of Michelson contrasts (algorithm 10) and, in particular, the ratios of items' contrasts to background (algorithms 11 and 12).

It should be noticed that some algorithms are mathematically identical, like the *Ratio of Relative Differences* (first entry in *Intermediate Ratio Computations)* and the *salmin* algorithm 5; these algorithms are listed in Table 1 but were not numbered (and, of course, not computed) twice. The same is true for algorithms that are not identical but make identical predictions when used to predict equal-salient targets; this is, for example, the case for the *Luminance Ratio* (no number) and the Weber contrast (algorithm 2), which reduce to the same equation of luminance ratios when used to predict test target luminance. Other algorithms may become identical in certain stimulus conditions, like algorithms 5 and 6 in minimum target conditions, in which the maximum and minimum are represented by distractor and background luminance settings. But these algorithms will differ in other stimulus conditions. An important example are the Weber and the Michelson contrasts, which make the same predictions for matches of similar targets (both DARK or BRIGHT and both maximum or minimum) but different predictions for matches of BRIGHT and DARK targets. Some algorithms, finally, may lead to identical predictions under the special restrictions of a particular test series, like algorithms 1, 2, and 3, which all had predicted equal target luminance for patterns in test series block *K* (where distractors were identical; see section A)[5].

The list of algorithms in Table 1 is rather long in spite of the simple stimuli used in this study. But although the stimulus patterns were simple, target salience can be, and obviously is, affected by various stimulus properties. For example, targets differ from distractors, but also from background, and both aspects were seen to contribute to their salience. Distractors, too, are salient, as they differ from background, and their *item salience* may also affect the resulting salience of targets. It is not *a priori* clear which of the different components can be ignored or must be added, subtracted, or weighted proportional to obtain the correct measure of salience in the experiments. Thus, while some algorithms in Table 1 only care of target and distractor luminance, others include the background as a parameter, and differences as well as ratios are used to combine target and distractor contrasts.

5 Small variations in MSD values from these algorithms are due to the fact that distractors were, in fact, not exactly identical. Since target and distractor luminance was slightly affected by stray-light on the monitor, the same luminance settings on a dark or bright background could slightly differ in luminance. These (measured) values were used in the computations.

Nonlinear luminance scales

While the first approach into the modeling of luminance-defined salience mechanisms was performed on linear luminance scales, certain tests (Exp. 11 and 12) had revealed matches that were better explained by salience computations on a nonlinear luminance transform. Perceptual sensation of certain luminance-based parameters, like brightness perception or discrimination threshold, are not linearly related to luminance but follow a power function of the form

$$\Delta\ sensation \sim \Delta\ luminance^{\,x}$$

(cf. Stevens, 1961). Experimental estimates of the exponent may go down to $x=0.33$ so that a linear change in perception might correspond to an up to 3-fold variation of luminance. To include these characteristics in analysis, all computations of Table 1 were repeated with values representing power transforms of the original luminance settings, with one of five different exponents; $x=1.0$ (the linear case already referred to above), *0.85*, *0.7*, *0.5*, and *0.33*. For that, background, distractor, and reference target luminance settings were first transformed with the according power function. The new values were used to predict, with the formulas of Table 1, the (transformed) test target values, which then were back-transformed into luminance (power functions with exponents *1/x;* values < 0 were ignored). All further computations and the evaluation of MSD were performed as already described in the General Methods section.

Conventions. For simplicity, I will use the nomenclature "power|x" to indicate that an analysis was based on the luminance scale after a power transform with exponent *x*; "power|1.0", for example, would indicate the linear case. When referring to certain algorithms, I will use the numbers listed in Table 1, but will put these numbers in rectangular brackets if the algorithm was applied to the power|0.33 transform. In predictions from algorithm 1, thus, targets had a constant luminance difference to distractors, whereas in predictions from algorithm [1] the power|0.33 transforms of these differences had been set constant.

The decision to use *all* algorithms of Table 1 for salience computations from power transforms was not driven by reasons of physical plausibility but rather by the easiness of table computations. After having replaced the luminance settings by their according power transforms, all computational procedures could be used with no need of further modifications. As to be expected, certain algorithms are "immune" against power transformations and produce the same predictions with power|x as with the linear power|1.0 luminance scales. For example, the power-transformed version of algorithm 2 (Weber contrast),

$$\frac{tg^x_{reference} - dis^x_{reference}}{dis^x_{reference}} = \frac{tg^x_{test} - dis^x_{test}}{dis^x_{test}} \quad ,$$

can be written as

$$\frac{tg^x_{reference}}{dis^x_{reference}} - 1 = \frac{tg^x_{test}}{dis^x_{test}} - 1 \quad \text{, and hence} \quad \frac{tg^x_{reference}}{dis^x_{reference}} = \frac{tg^x_{test}}{dis^x_{test}}$$

which gives the same result of tg_{test} for all exponents *x*. The same is true for, e.g., algorithm 3 (Michelson contrast).

Additional Algorithms (tested but not reported here)

A number of algorithms not listed in Table 1 were tested but not included in the final analysis.

Hybrid algorithms. As already discussed, targets with low item contrast are barely visible and cannot be salient, even when they strongly differ from distractors. This was taken care of in an intermediate analysis with several "hybrid" algorithms. Predictions from these algorithms were based on separate computations, one calculating the presumed target-to-distractor (discrimination) salience, the other the presumed target-to-background (item) salience. Hybrid algorithms were not generated for algorithms that did already include combinations of both salience measures, like all algorithms above number 10. The two salience measures were then combined by taking the minimum of both. Other combinations (like multiplication, division, or different weightings of the two measures) were not calculated. Hybrid algorithms of this sort were only applied to targets in *minimum target configurations*; since only in these configurations could low item salience block the higher discrimination salience. Some of these algorithms produced quite good predictions of the data, but these were generally not better than the best fits obtained with the "standard" algorithms in Table 1. Therefore, the results from this sort of hybrid algorithms are not reported here.

A different way to combine discrimination and item salience measures might have been the evaluation of threshold values to maintain the target's visibility (as discussed in Fig. 15). This way was however not implemented, for two reasons. First, subjects had apparently used *different* thresholds in their adjustments (see Fig. 15b) and further experiments would be necessary to measure these thresholds in order to put the computations on an experimentally valid ground. Second, all these variations could explain only small deviations in a very *small proportion* of experimental data (mainly a few minimum targets in very small distractor-to-background luminance window). At the present stage of the project it seemed more important to clarify the general outlines of salience computation from luminance than to make exact predictions of small modifications in certain stimulus combinations.

In a later stage of the study, a different *new hybrid model* was tested, which computed various *averages* of predictions from algorithms 1 *(constant addition)*, 2 *(constant ratio)*, and 5 *(salmin)*. The model seemed particularly promising in predicting salience matches of BRIGHT maximum targets and will be introduced below.

Root of Mean Squares. An algorithm that, on a first glance, should have been useful to measure target conspicuity in different luminance settings is the RMS (root of mean squares) which provides a reliable measure of contrast in textures and random dot patterns (Moulden, Kingdom, & Gatley, 1990). However, with the simple stimuli of the present study with only two luminance levels beside the targets, the RMS cannot generate new predictions over those already listed in Table 1. Another objection against a general testing of this algorithm is the wide spacing of items in the present study, which does not generate the percept of textures or random dot patterns, for which the RMS measure was suited best (Moulden et al., 1990).

Reflection and Self-Luminosity. Finally, a different line of modeling should be mentioned that would be based on an interpretation of luminance settings as reflection and self-luminosity components. In principle, the luminance of all items and the background in the tested patterns could have represented a combination of reflection and self-luminosity, and all six parameters might have changed independently between the reference and test patterns. To allow for any useful predictions from this model, we would have to make assumptions about item and background identities and about illumination changes between the patterns. Plausible (but arbitrary) assumptions might, for example, be that backgrounds were purely reflective and any (inconsistent) variations of distractor and target luminance resulted from additional luminosity components of the items. But further restrictions are still required to predict test target luminance.

With a few such restrictions the model was originally included in the computations for Experiments 1-4 (section A). But the model was later removed from analysis as it could not be applied to all stimulus combinations tested. The plausible assumptions on which the model was originally based cannot account for targets with different luminance polarity (DARK versus BRIGHT) or for targets with different rankings (minimum versus maximum) in reference and test patterns. Targets that are dimmer than background or distractors cannot be transformed, by a pure illumination change, into test targets that are brighter than background or distractors.

Computations

Each algorithm of Table 1 was used to predict the luminance of salience-matched targets in all test conditions of the study. With the plausible assumption that

$$sal_{test\ target} = sal_{reference\ target}$$

(equation 1) for a perfect salience match, the luminance settings of each test condition were entered into the formulas and solved for tg_{test}. With some formulas, this leads to quadratic equations with two solutions, of which only the physically possible one (luminance cannot be negative) or that closer to the experimental data was taken. The quality of each individual prediction was measured as the mean squared deviation, *MSD,* of predicted from measured data (averaged over all subjects). MSD values were computed separately for the different test series blocks and also for "totals" of all tested stimuli. Dense and wide blob arrangements were analyzed both separately and in combination.

For combinations of different targets, as studied in section B, two different procedures were used for computational predictions. In one procedure it was assumed that the salience of the different types of reference and test targets was obtained from *different* salience mechanisms. Predictions were thus achieved from *mixed combinations* of two algorithms, one to compute the presumed salience of the reference target and the other to predict from this salience value the expected test target luminance setting. To restrict the number of combinations, only the five best algorithms of either target type (as listed in Table 2, column 5) were used for predictions from *mixed combinations*. For the other procedure it was assumed that the salience matching of different targets is based on a common mechanism (which might be different from that one optimal for pure salience matches of either target type). Thus, reference target salience computation and test target luminance prediction were achieved from the same *unique algorithm* in a similar procedure as that used for data from section A.

Intrinsic Correlations. To find out how correlated, or uncorrelated, the various algorithms behave when applied to the data, a virtual set of luminance data was tested, in which background, distractor, and reference target luminance settings were systematically varied over the experimental range and test target luminance was computed using the algorithms of Table 1. The correlation coefficient of different such computations was computed.

Outliers. To avoid extreme (and non-realistic) deviations, predictions that fell outside the experimental luminance range were ignored and not included in the analysis. In these cases, however, the MSD was based on a smaller number of data points than tested and a smaller number of data than evaluated in other computations, which could have biased the results. To overcome this problem, the MSD computation (for a given algorithm and block of test series) was defined as invalid and excluded from analysis, if 10**%** or more of the data points were ignored.

A different sort of "outliers" was data for which a perfect match could not be obtained experimentally, for example, if the matches had fallen outside the luminance range of the monitor. Observers could not reject such matches but always had to find the best approach. Most of these mismatches were obvious and data points were then removed from analysis. But even if a few of such mismatches had not been removed, the resulting error should be small. If such data points had been correctly predicted, the resulting values should have fallen outside the monitor limits and data points hence been excluded from MSD computation.

Results and Discussion

An overview of the best fitting algorithms for each experiment is given in Tables 2, 2H, 3, 3A and 4. Some of these fits were already plotted in the previous figures. As we will see, there were notable differences in the quality of predictions. In some test series, one algorithm was clearly better in predicting the experimental data than others; in other test series, several algorithms might have produced comparable and sometimes rather good fits. In particular the matches of DARK targets in *minimum* configurations were very closely predicted by quite a few algorithms (cf. Table 2 C1).

Intrinsic correlations

Given the limited data set to which the algorithms were applied (six parameter values, the independent variation of which is often reduced by general restrictions from target type and configuration) and the similar "essentials" in the computations (most algorithms compute, in various combinations, the target-to-distractor difference) it is not surprising that many algorithms make similar predictions in certain tests. In fact, when the available luminance range is systematically scanned with background, distractor, and reference target settings and test targets are predicted from these settings, the data show a high intrinsic correlation. Over all possible pairs of algorithms, the averaged correlation coefficient is $r = 0.67$, without the identity cases. For example, the correlation coefficient of algorithms 1 (constant addition) and 2 (constant ratio), averaged over all possible test conditions, is $r = 0.86$; it varies between 0.64 (for matches of DARK minimum and DARK maximum targets) and 0.94 (for matches of BRIGHT targets in either configuration). Only two algorithms, algorithms 15 and 16, produced data that were, on average, less strongly correlated (and for certain target matches even anti-correlated) with the predictions from other algorithms (mean correlation coefficients, $r = 0.33$ and $r = 0.30$); they both do not depend on distractor luminance. But even these algorithms were more strongly correlated with other algorithms in certain tasks. Thus, it is unavoidable that the possibly correct salience mechanism shares some of its features with other algorithms of Table 1.

Strong intrinsic correlations also exist when algorithms are applied to either the linear luminance scales or their power transforms. The averaged correlation factor between these cases is r = 0.94 when averaged over all algorithms (without the identity cases for which computations do not change with the exponent of power transforms). But correlation factors can be smaller, and computations more distinct, when the investigated luminance range is large and luminance values differ considerably from their power transforms.

I. Matches of Similar Targets

Hints for reading: This section analyzes the computational predictions for Experiments 1-4; which algorithms are good and which ones fail. Fig.41 shows improved fits of BRIGHT target data when *averages* of certain predictions rather than the individual predictions themselves are looked at. Otherwise there is nothing new here if you are not interested in details.

In the first major computational analysis, we attempt to predict performance of equal-salience matches in Experiments 1-4 of section A and Experiment 7 and 8 of section B.

MSD variations with different algorithms

Figure 37 shows the MSD values obtained with different test series on *DARK maximum targets*; the smaller the MSD value for an algorithm, the better is its fit to the experimental data. Algorithms are sorted for increasing MSD values in the *Total* test sample on the left; the same order is used in the presentations of the different test series towards the right. With *test series block L*, quite a few algorithms produced more than 10% invalid predictions (values outside the available luminance range) so that MSD values were not taken.

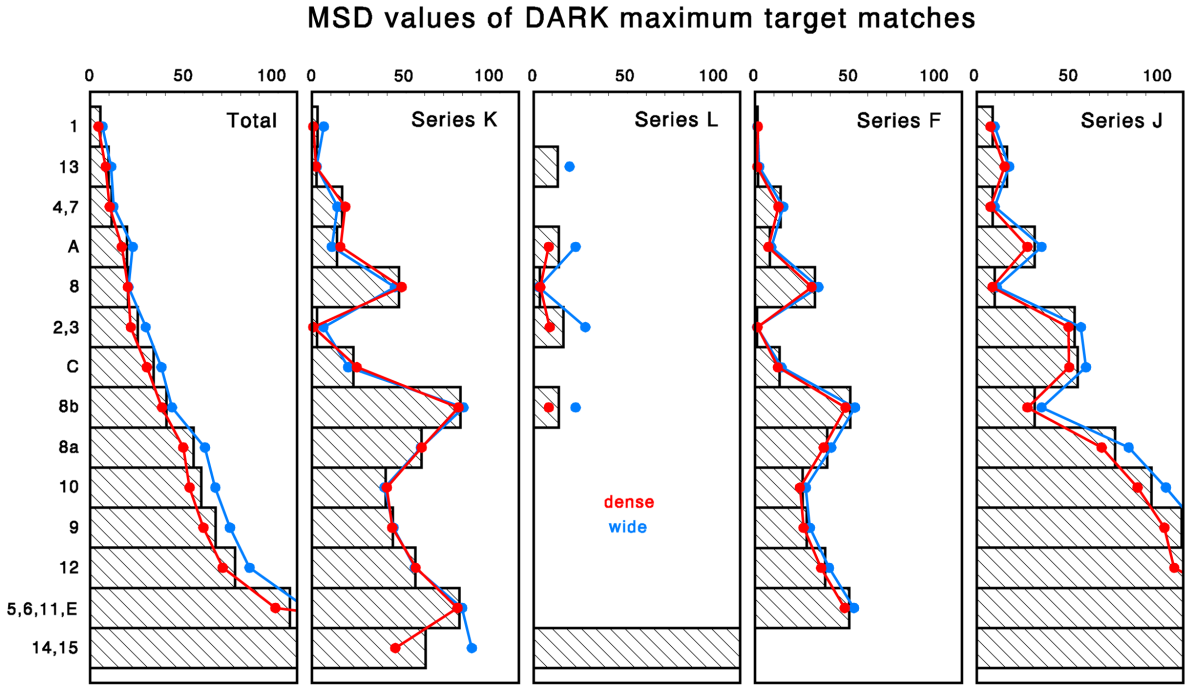


Figure 37. *Distributions of MSD (mean squared deviation) values for various predictions of DARK target matches in the different test series of section A.* Algorithms (see Table 1) are listed on the left and are sorted for increasing MSD values in the Total test sample. MSD values of the other test series are shown in the same sequence. Different algorithms listed in one line were mathematically equivalent in these tests or made identical predictions. Histograms plot MSD values for *all* tests in a given test series; red and blue curves those of the sub-samples obtained with dense and wide blob configurations, respectively. All values are plotted towards the right (see scales on top). Missing data indicate that the algorithms had produced too many invalid predictions, i.e. predictions outside the available luminance range of the monitor (see text). By and large the general courses of distributions are similar for different test series, except for local “sinks” in some distributions. For example, algorithm 2 (identical with algorithm 3 for these matches) produced particularly small MSD values in test series *K* and *F*, but not in series *J* or the *Total* sample. The overall smallest MSD values for DARK maximum targets were obtained for predictions from algorithm 1.

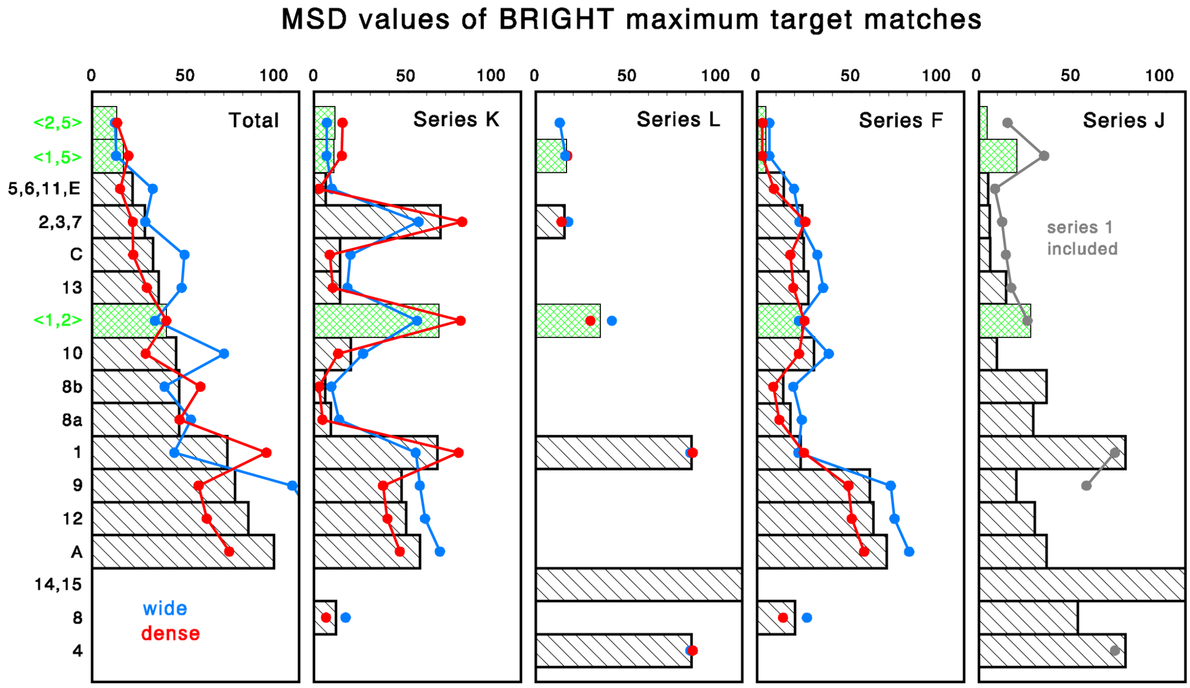


Figure 38. *Distributions of MSD values for predictions of BRIGHT maximum target matches in section A.* Presentation as in Fig.37, with two modifications. In addition to the algorithms listed in Table 1, also data from three hybrid algorithms are included (green), which represent averages of predictions from three original algorithms (see text). In test series J where data from wide and dense blob matches had been pooled (see Fig.20), the superimposed data curve (gray circles) represents MSD values if condition 1 (label 1 in Figs.20 and 21) was *included* in analysis (in the histogram data it was not). The overall smallest MSD values were obtained from the hybrid algorithms <2,5> and <1,5> and from algorithms 5 and 2 (equivalent to algorithms 6, 11, E and 3, 7, respectively).

While some algorithms produced large MSD values and hence failed to predict the experimental data (e.g., algorithms 5 and 14), other algorithms generated very small MSD values and thus closely described the observers' performance in the salience matches. In the *Total* sample, the smallest MSD value was 5.4[6] obtained with algorithm 1 (Table 2, A5). Certain algorithms were perfect for one particular test series but less perfect for another series. For example, algorithm 8 made particularly good predictions in test series block *L* and *J* but was poorer in test series blocks *K* and *F* where other algorithms produced much smaller MSD values. In the *Totals* computed over all test conditions, algorithms 1, 13, 4=7, A, and 8 (in this order) were best (Table 2, A5). The constant-ratio principle (Weber contrast; algorithm 2), on the other hand, which should be very useful for salience estimates because of its insensitivity to illumination changes, was poorer than any of these. It only made good predictions in test series blocks *K* and *F*, but that might be due to the special stimulus conditions in these tests (identical distractors in reference and test patterns), which made this algorithm indistinguishable from algorithm 1. Beside such variations, however, the general ranking of MSD values was fairly similar in the different test series. Certain algorithms were good, others bad, in *all* test series. Note that the *salmin* algorithm 5 (which had fitted the salience matches of several other target types; cf. Figs. 3a, 7) was generally very poor with DARK maximum targets. There was a small variation between matches in dense and wide blob configurations (blue and red MSD curves, respectively, in Fig. 37); predictions were generally a little better for data obtained in dense configurations.

In the MSD distribution for *BRIGHT maximum targets* (Fig. 38) there were more invalid computations, and hence more missing data entries. Across these holes, however,

6 For the physical unit of MSD values, see Footnote 1.

the trends are similar for the different test series. Distributions are again sorted for increasing MSD values in the *Total* test sample, but except for a few outliers the ranking is similar in the other test series blocks. Only series *K* shows notable deviations; there are small and large MSD values all over the sequence. One particular high MSD value was obtained for algorithm 2 (and equivalent algorithms 3 and 7) which made better predictions in the other test series. Performances in wide and dense blob configurations (blue and red data points) again are similar although more variable than with DARK maximum targets. An obvious difference compared to DARK maximum targets (Fig. 37) is the larger MSD values of even the best predictions (smallest *black* histogram bars). The best algorithm 5 (equivalent to algorithms 6, 11, and E) for BRIGHT targets produced, in the *Total* sample, an MSD value of 21.4, which is about 4-fold the best MSD value 5.4 obtained for DARK maximum targets (Table 2, E5 vs. A5).

New hybrid algorithms

This poorer performance of even the best algorithms with BRIGHT targets has set off the search for a new algorithm that might better predict the experimental performance. Several experiments have suggested that matches might sometimes fell in the middle of two predictions, so as if observers would match target salience according to two models and then chose the mean of both (cf. Figs. 3b, 11 and 12). To test if such an averaging process might have also been used in the salience matches of BRIGHT targets, three new algorithms were tested defined as algebraic means of two (of three) standard algorithms. Algorithm <1,2> computed the means of predictions from algorithms 1 and 2, algorithm <1,5> the means of predictions from algorithms 1 and 5, and algorithm <2,5> the means of predictions from algorithms 2 and 5.

The MSD values of these averaged predictions are plotted in green into Figure 38. An overview is also given in Table 2H. Two of these new algorithms did indeed improve the predictions of experimental data in quite a few but not all test series. Algorithms <2,5> and <1,5> provided the best fits of the experimental data in the *Total* sample and in test series *F*, but were not better than the best predictions of other algorithms, e.g. algorithm 5, in test series *K, L,* and *J*. Algorithm <1,2>, on the other hand, generally made poor predictions of the experimental data.

There are several aspects in these data, beyond a simple listing of performance ranking, that are worth further analysis. To understand and evaluate the differences between test series it is important to look into predictions in detail. It might also be helpful to compare the MSD distributions of different target types and, in particular, their variations associated with linear or power-transformed luminance scales.

Fits and misfits in predicted performances

Figures 39-41 show predictions from different algorithms in a number of tests; the examples illustrate various aspects of the fits and deviations for targets in maximum configurations. A detailed inspection of predictions with targets in minimum configurations is here not made, since predictions were generally quite good for these targets.

DARK maximum targets

As was already seen in Figure 37, the quality of predictions from the same algorithms varied between test series. This is visualized in Figure 39 where predictions (colored curves) are directly compared with the experimental data (black squares). For example, predictions from algorithm 8 (purple) closely meet the experimental data in test series *L* and *J* but strongly deviate from data points in test series *K* (label 1) and *F* (label 2). Obviously, the different backgrounds in the latter two series (not shown here, but see Figs. 3 and 9) affect predictions more strongly than was the case in the matches themselves. Instead, in test series *K* and *F* algorithms 1 (lighter blue) and 2 (yellow; partly hidden behind the lighter blue line) produce almost perfect fits, but predictions from these algorithms, deviate more strongly, at least locally, from the data in series *L* and *J*. Only algorithm 13 made predictions that are fairly close to the data in all test series, but these predictions are consistently less perfect than the best predictions from either algorithms 1, 2, or 8. In test series *J*, we see an increasing deviation of predictions from the data from the left-hand curve sections (label 4) towards the right-hand sections (label 5). Remember that reference distractor levels varied between these curve sections whereas background settings had remained constant over all tests (cf. Fig. 17). On the left-hand side of series *J* in Figure 39d (curve section at label 4), data represent identity matches, and expectedly all algorithms accurately predict this performance. But with increasing luminance levels of the reference

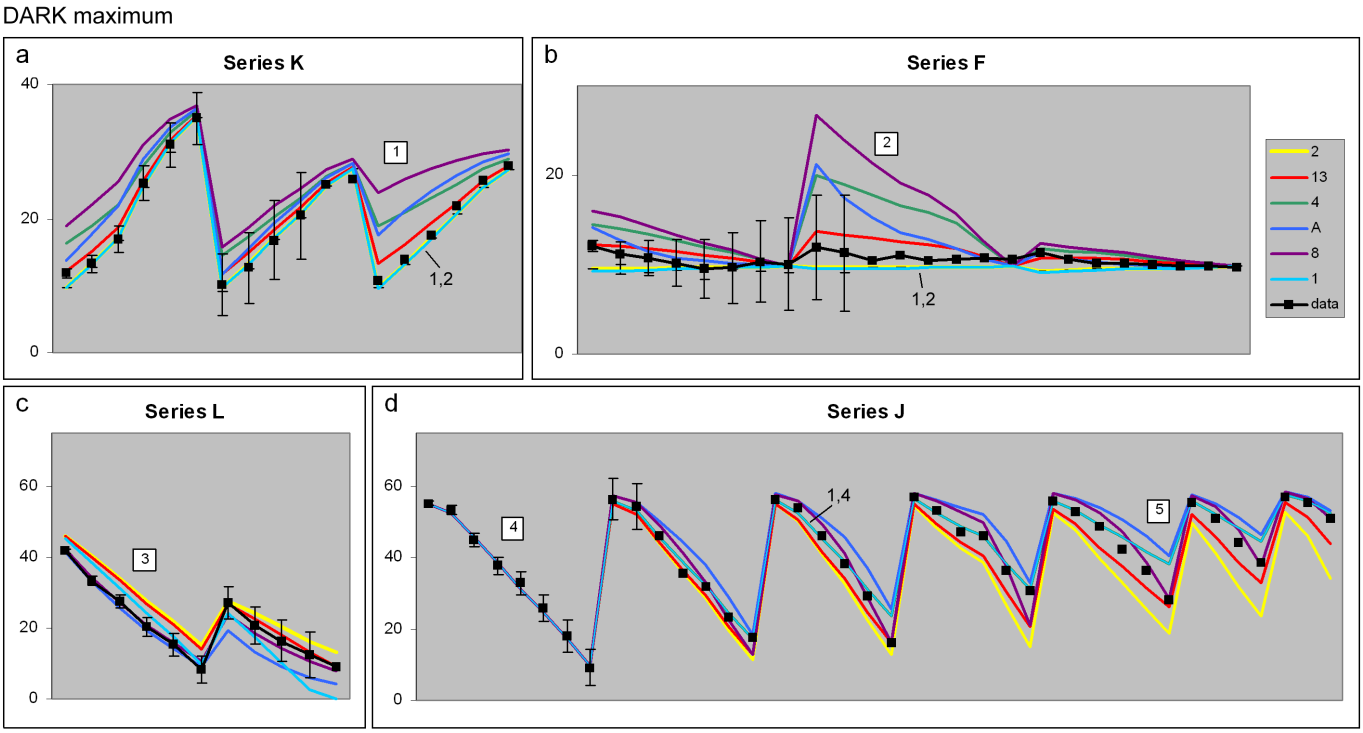


Figure 39. *Best predictions of DARK maximum target matches.* Plots show sequences of matching data obtained in section A (black squares) together with best predictions (colored curves) obtained from various computations in section C. **a.-d.** Different test series blocks as indicated; note that the data from different series within each block, which were plotted separately in the previous figures, are here (and in the subsequent figures) shown as continuous sequences. Curves listed but not visible in a certain plot are hidden behind another, similar curve. Framed labels are referred to in text.

distractors (curve sections towards the right), reasonably good predictions were only maintained from algorithms 1=4, 8, and 13, while predictions from algorithms 2 and A develop to become the worst ones of this collection. Over the *Total* data sample, however, algorithm 1 turns out to be the best (cf. Table 2, A5), although it is beaten in some test series by other algorithms. Algorithm 8, on the other hand, which predicts quite accurately the experimental data in test series block *L*, is too sensitive to background variations and thus fails in other tests. To get a feeling of how deviations are reflected in the MSD values, the accuracy of individual fits in Figure 39 should be compared with the according values in Table 2, row A. But note that the MSD is computed from all test conditions of a series; therefore, even large deviations in a few test conditions may still result in an acceptable MSD if all other test conditions are met. Also note the different *y* scales in Figure 39.

BRIGHT maximum targets.
Similar observations can be made with BRIGHT targets (Fig. 40). Algorithm 2 (red lines), for example, generated the best predictions in test series blocks *LX* (label 6) and *J* (label 10) but strongly failed in test series blocks *K* (label 1) and *F* (label 7). Note that problems occurred when BRIGHT test targets were to be predicted from dimmer reference targets; several algorithms then predicted luminance settings that exceeded the monitor settings in experiment. This is seen in the examples of test series *L (labels 2 and 3)* and *J* (at label 8). Even when these data points are excluded from analysis, the remaining predictions might falsely be considered a too good fit (small MSD), despite the fact that many data points could not be fitted at all. This was the reason to introduce a 10% rule according to which MSD values were not computed when 10% or more of the data points in a sample fell outside the monitor range and had to be ignored. After

BRIGHT maximum

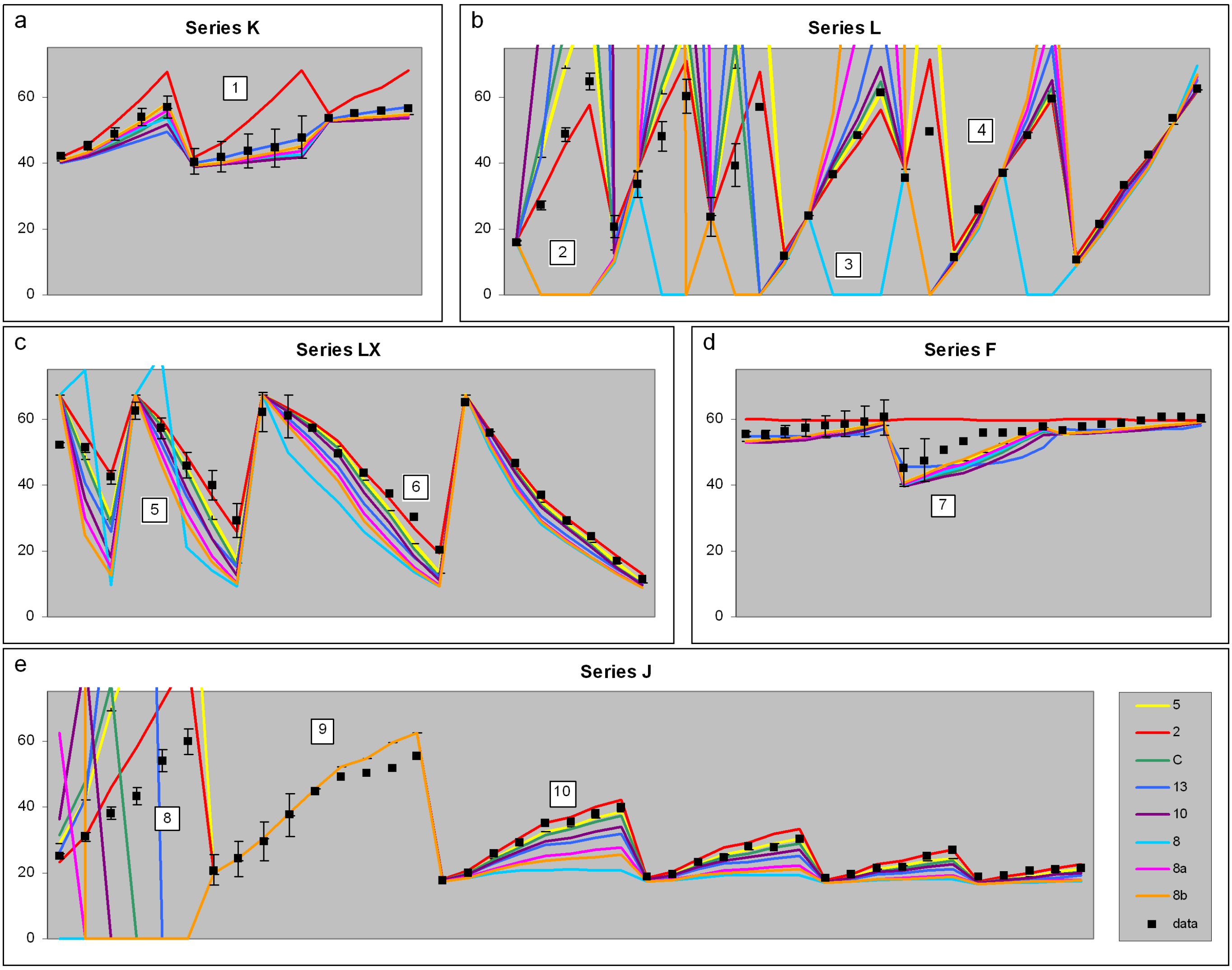


Figure 40. *Best predictions of BRIGHT maximum target matches.* **a.-e.** Data sequences from different test series blocks, as indicated. Presentation as in Fig.39. For details see text; framed labels will be referred to there. Zero values (labels 2, 3, 8) are from invalid predictions.

applying this rule, only MSD values from algorithms 1, 2, 4 and 14 remain valid in series *L* (cf. Fig. 37), of which only algorithm 2 produces sufficiently good predictions to be plotted in Figure 40 (red lines). But note that exactly this remaining algorithm 2 has made rather poor predictions in test series block *K*. *Vice versa*, the best algorithms for series *K*, algorithms 8b, 5, 8a, 8, and 13 (cf. Table 2, E1), all generated too many missing data points in test series *L*. The amount of invalid predictions was much smaller in test series block *LX* (labels 5 and 6) where test targets were dimmer than reference targets and most predictions did not exceed the available luminance range. To overcome the problem of invalid predictions in the left-most data points of test series block *J* (where test targets had to be adjusted brighter than reference targets; cf. Figs. 20c and 21), MSD analysis was restricted to test conditions with test targets equal or dimmer than reference targets (all conditions from the second curve section on in Fig. 40, label 9); the “problematic” matches (label 8) were excluded in MSD computations. Since subjects had apparently switched to a different algorithm in these tests (cf. Fig. 21), it is reasonable to not analyze the data in one common sample.

The interpretation of performance in the other test series is straight forward. In test series *F*, algorithm 2 is not affected by background variations and hence cannot predict the experimental performance, which is clearly affected by background settings. All other algorithms follow the data variations, the strongest deviations being produced from algorithm 13. Thus, Figure 40 reveals several best fitting algorithms, but none of them is good in every test series. Algorithm 2 is rather good in test series blocks *L, LX*, and *J*, but poor in test series *K* and *F*. Algorithms 5 and C are good in most test series (in series *J* even better than algorithm 2) but produce a large number of predictions outside the available luminance range in series *L*. Altogether, the best MSD values of the *Total* sample are therefore larger with BRIGHT than with DARK maximum targets (Table 2, E5 vs. A5).

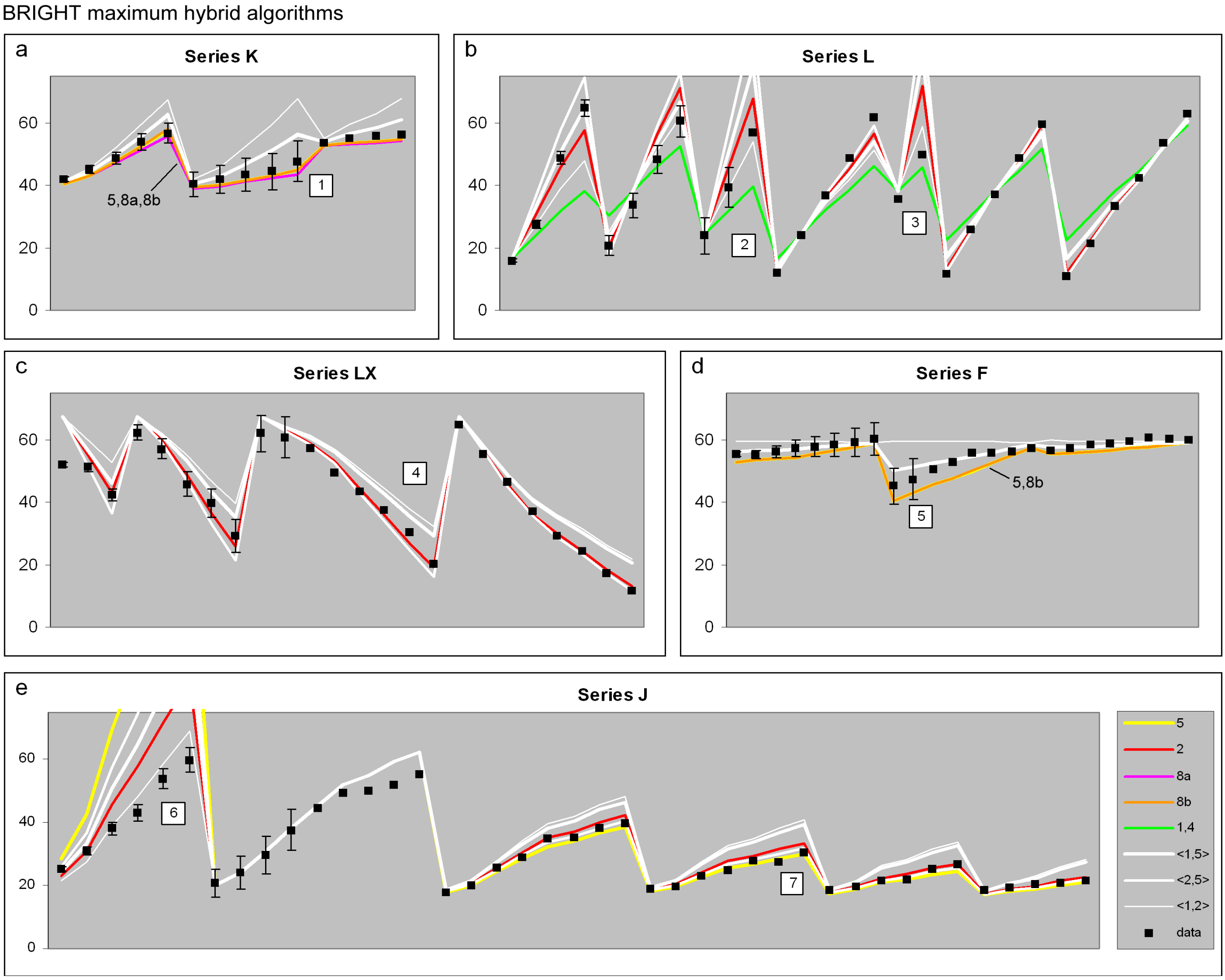


Figure 41. *Fit improvements in Fig.40 by hybrid algorithms.* **a.-e.** Data sequences and best predictions from Fig.40 are shown together with predictions from the three hybrid algorithms <1,5>, <2,5>, <1,2> (white). For discussion, see text.

The partially improved fits of the *new hybrid algorithms* <1,2>, <1,5>, and <2,5> are shown in Figure 41 as white curves. While the thinnest white lines (predictions from algorithm <1,2>) sometimes run far off the experimental data (e.g., at labels 1, 4, 5 and 7, but not at labels 2, 3, and 6), the middle-thick lines (predictions from algorithm <2,5>) could get very close (e.g., at label 1, here the middle-thick white curve is partially hidden by the colored lines; and in test series blocks *LX, F,* and *J*, again partly hidden) although often not closer to the data than other lines, i.e. predictions from other algorithms. A strong improvement is seen in test series block *F*, where predictions from the remaining hybrid algorithms <1,5> (thick white line) and <2,5> (middle-thick line; partly hidden by predictions from algorithms 5, yellow, and 8b, orange) both fall close to the data, and a gradual improvement in test series block *J*, where predictions from algorithm <2,5> (middle-thick white) fall right in between the already good predictions from algorithms 2 (red) and 5 (yellow), as expected from the definition of the hybrid algorithm. Altogether, thus, the introduction of new hybrid algorithms had gradually improved the fits of experimental performance (Table 2H).

To illustrate the improvement with test series block *F*, part of Figure 9 is re-plotted in Figure 42 now showing the original matching data (blue and red symbols) with

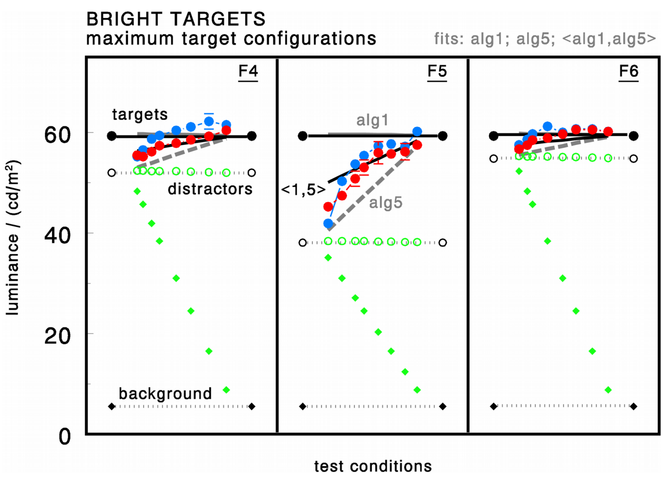


Figure 42. *Fit improvements by hybrid algorithms in Fig.9c.* Data plots from Fig.9c (constant BRIGHT maximum targets on variable background) are reproduced; predictions from the hybrid algorithm <1,5> are shown together with the predictions of both components. The averages fit the data better than either algorithm alone.

predictions from the constant-addition principle, algorithm 1 (continuous gray lines), the *salmin* algorithm 5 (dashed gray lines) and from the hybrid algorithm <1,5> (black lines). Except for test targets that were close to distractor luminance (green circles), matches are better described by the means of these two predictions than by the predictions themselves.

MSD distributions

Figure 43 gives an overview of the MSD distributions for the *Totals* of all different target types. The data are individually sorted for magnitude; hence the sequences of algorithms differ between the graphs. The upper two histograms replicate the *Total* sample plots of Figures 37 and 38, but red symbols now represent different data. The lower two curves plot the MSD distributions of target matches in minimum target configurations.

There are notable differences between these distributions. The MSD values for *maximum targets* are, on average, larger than the MSD values for minimum targets. They start at larger values and increase to finally larger values on the right-hand side of the histograms. In contrast, the distributions for *minimum targets* show many very low MSD values on the left-hand side and only few larger values towards the right. That is, more algorithms made predictions outside the monitor range for minimum target conditions and then had to be excluded from analysis. But of the remaining algorithms also many more fitted the salience matches of minimum targets much better than those of maximum targets.

This difference cannot be explained by differences in the luminance ranges covered by these tests. Although targets in minimum configurations had to be adjusted between two luminance settings (the background and distractor levels), the overall luminance ranges were similar in all tests, as can be seen in Figures 3, 5, and 7. The fact that the upper and lower luminance limits were set in the matches did not *a priori* restrict the range of *predicted* test targets. In fact, quite a few algorithms predicted targets outside this range (and sometimes even outside the limits of the monitor) and algorithms had to be excluded from MSD evaluation if this happened too often in a sample. Note that the MSD distributions in Figure 43 include only six algorithms (histogram bars) for DARK minimum targets but fourteen for DARK maximum targets.

A general problem in the prediction of test target luminance are formulas that include ratios of higher to lower luminance settings, like the Weber contrast (algorithm 2) for BRIGHT maximum targets. When the distractor luminance in the reference pattern is low, ratios can become very large, and the prediction of an equal-salient target on an only slightly increased distractor luminance in the test pattern may quickly exceed the available luminance range. The problem will not occur with DARK maximum targets, as their luminance ratio to distractors remains below 1 and cannot become smaller than 0. This difference in the computation of luminance ratios between dark and bright stimuli is one reason why more MSD values had to be excluded with BRIGHT maximum targets (11 histogram bars in Fig. 43, without the later added hybrid algorithms) than with DARK maximum targets (14 bars). For minimum targets, the difference is reversed; there are more valid MSD computations for BRIGHT (10 bars) than for DARK minimum targets (6 bars). This is explained by the reversed target-to-distractor polarity in these patterns. Although in, e.g., DARK minimum target configurations all items are dimmer than the background, the targets are *brighter* than the distractors. In BRIGHT minimum target configurations it is the other way around; targets are dimmer than distractors, even though they all are brighter than the background.

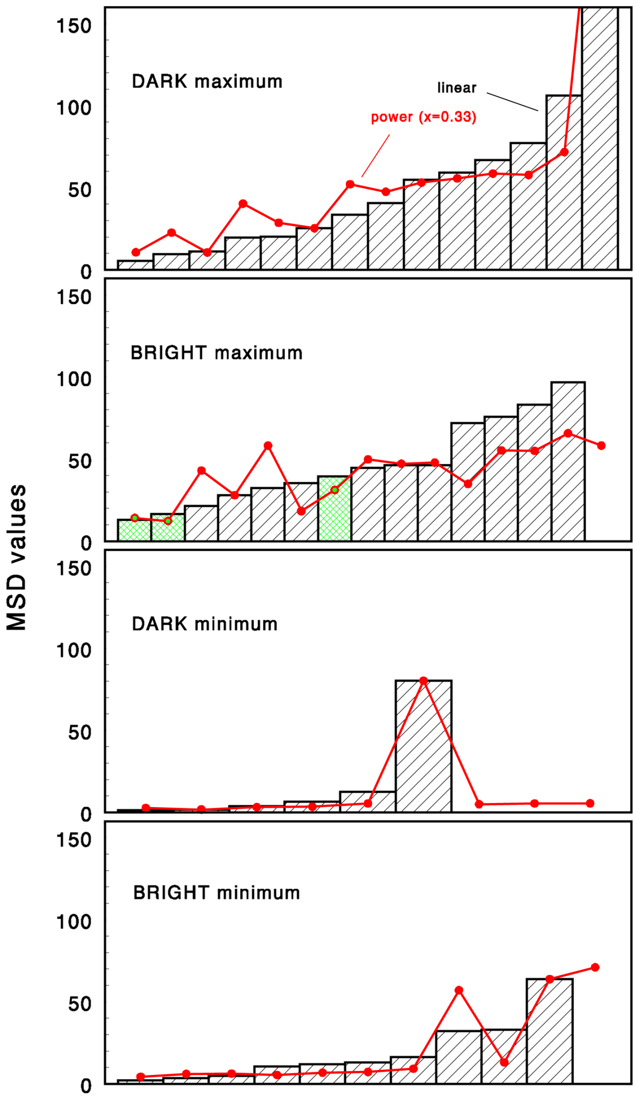


Figure 43. *MSD distributions of the four target types tested in section A.* Histograms plot the values obtained for the Total sample in increasing order; green bars (in the distribution for BRIGHT maximum targets) are from the three hybrid algorithms. Superimposed red curves show the MSD values obtained for the same algorithms when applied to the power-transformed luminance scale (exponent x=*0.33*). There are only minor differences between the two analyses. Algorithms with valid data in the power|0.33 analysis (red circles) but no valid data in linear analysis (histogram bars) were added to the right.

Predictions from nonlinear luminance scales

An interesting aspect is the effect of luminance scale transformations upon the quality of target predictions. The analysis presented so far (bar histograms in Fig. 43) was based on predictions from linear luminance scales. The red curves show the same analysis based on nonlinear luminance scales obtained from the power|0.33 transform, which was the strongest tested deviation from linearity. The MSD values are plotted in the same order as for the linear cases; that is, superimposed bars and data points refer to the same algorithm. Algorithms that had to be excluded from linear analysis (because there were too many predictions outside the accepted luminance range) but did generate reliable MSD values in the nonlinear analysis are added on the right-hand sides of the histograms without accompanying histogram bars.

As Figure 43 shows, the use of power|0.33 transforms did, in general, not affect the quality of predictions. For a few algorithms, the predictions are slightly improved (smaller MSD values), for others the previous predictions from the linear power|1.0 luminance scale are better. In general, predictions from both scales were correlated and were either both high or both low for a given algorithm. The two MSD distributions for a given target type were

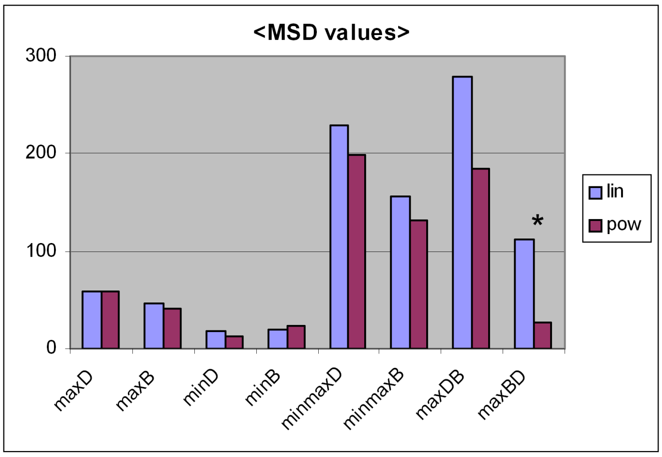


Figure 44. *Mean MSD values of the various matches analyzed in section C.* In the abbreviations below, *D* and *B* stand for DARK and BRIGHT target matches and *max* and *min* for targets in maximum or minimum configuration; the code is not repeated for matches of similar targets. Differences between linear *("lin")* and power|0.33 analyses *("pow")* are much larger for the cross-polarity matches "maxDB" and "maxBD" (but only for the latter one significant; $p<0.05$).

similar and best MSD values were not notably reduced when the linear luminance scale was replaced by its power transform. The means of each distribution do not show significant differences between linear and nonlinear luminance scales (Fig. 44).

In summary, thus, the thorough analysis of salience matches of similar targets with a large variety of algorithms did not disclose the one unique salience mechanism that would explain all experimental data, but revealed a number of algorithms that could perfectly predict the data in some test conditions but not in others. The best algorithms for the different target types are those with particularly small MSD values, but small differences between these values must not be stressed, as predictions from various algorithms have shown both fits and failures in different data samples. With BRIGHT maximum targets, the overall quality of predictions was slightly reduced compared to that with DARK maximum targets, when the magnitude of MSD values is looked at. Here an improvement was achieved by predicting salience matches not from single algorithms but from averages of two different algorithms, suggesting that salience is perhaps less exclusively represented by a single mechanism than one might have assumed. Perhaps one of the most important findings of section C so far is the observation that fit performance did generally not improve when analysis was based on power-transformed (power|0.33) instead of linear luminance scales, suggesting that power transforms are only little important in salience matches of *similar* target types. As we will see below, this finding is different in salience matches of *different* target types and, in particular, of targets with different luminance polarity to background.

Conclusions from computations part I:
Equal-salience matches of *similar targets* can be closely predicted but best algorithms vary with target type, target configuration, and test sample. – Predictions of BRIGHT targets are often improved when the means of the salmin and constant-addition or constant-ratio predictions are taken rather than any of these predictions itself. – Power-based computations are not better than linear ones.

II. Matches of Different Target Types

Hints for reading: Still alive? I' am afraid it's getting even worse now. The following section reports predictions of experiments in section B, with *different* targets: many details in detailed analyses – skip if you are merely interested in an overview. Take-home messages are: mixed combinations are, in general, not better than unique algorithms, even if these could not explain the salience estimates of the involved targets in other matches. Cross-polarities matches (but not equal-polarity matches) are best predicted from power-transformed luminance scales.

In the following analyses, algorithms of Table 1 will be used to predict the equal-salience matches *in section B*. We will first look at matches of similar targets in minimum vs. maximum configurations (Exp. 9-10, 12 and 13) and finally at matches of targets at opposite luminance polarities to background (Exp. 11 and 12).

Selected Algorithms: Mixed vs. Unique

When trying to predict the matches of *different* target types and target configurations, it seems logical to start with a combination of those algorithms that made the best predictions for exactly these target types and configurations. For example, in the matches of DARK minimum (reference) and DARK maximum (test) targets (Experiment 9), the salience of the reference target should be best predicted from algorithms 10, A, or C (cf. Table 2, C5) and that of the test target from algorithms 1 and eventually algorithms 13 or 4 (Table 2, A5). It would thus be plausible to start the analysis with predictions from *mixed combinations* of these algorithms.

Mixed Combinations

To reduce the number of computations, combinations were restricted to the five best algorithms for either target (as listed in Table 2, column 5) plus additional algorithms with best performance on the power-transformed luminance scales (also listed in Table 2, column 5) if these were not yet included. In the example of DARK minimum and maximum target matches, this resulted in six algorithms from Table 2, C/D5 (algorithms 10, A, C, 12, 5, and 8) to be combined with six algorithms from Table 2, A/B5 (algorithms 1, 13, 4, A, 8, and 2), hence target predictions from 36 possible combinations. Similar selections were made for the other target combinations tested in section B. Many of these predictions were rather poor (cf. Table 3 and 3a) and do not nearly reflect the good performance of the individual algorithms in equal-target matches. With *DARK targets*, for example, the combination of the best algorithms 10 and 1 (Table 2, C5 and A5) produced predictions with MSD values of 219.9 *(Series J)* or 194.4 *(Totals)*, which did not even make it into the lists of the five best mixed combinations in Table 3, A1 and A7. Only the *sample2* data from wide blob arrangements were about fitted by this combination (Table 3a, A1c). On the other hand, combinations of originally less perfect algorithms, like algorithm 5 for DARK minimum and algorithm A for DARK maximum targets, produced the best fit of the full experimental data set in Series J (Table 3, A1) and the second-best fit in two sub-samples of the data (Table 3a, A1a and b); according to the MSD values, however, none of these fits is very good.

It is important to point to a general pitfall of *mixed combinations*. As equations *(1)* and *(1a)* show, formulas in Table 1 are connected to salience with a proportionality factor *k*. Only if the salience of both targets is computed from the same algorithm (equation 1a), that factor will cancel out and can be omitted. If different algorithms are combined, as in the *mixed combinations*, likely two different factors, k_1 and k_2, must be assumed. This is particularly obvious for the algorithms 1 and 8, 8a, 8b, which already differ in their physical dimensions from all other algorithms. Most algorithms calculate luminance divided by luminance, with no physical unit of salience, but algorithm 1 computes luminance (unit cd/m^2) and algorithms 8-8b compute luminance divided by squared luminance (unit $1/(cd/m^2)$). Just from these differences *mixed combinations* with these algorithms should likely be invalid. But even if the according proportionality factors would be given a unit to make the equations physically correct, their values are not known. While it should be simple to compensate for the different factors k_1 and k_2 in the formulas, it is not at all obvious how to derive these factors from the equal-salience matches. I have tested two (arbitrary) procedures. (i) For each of the two combined algorithms I have computed the *maximal* salience value in a test series; these maxima were then equated to give the relative weighting of k_1/k_2 for the predictions. (ii) I made the same salience computations for *each data point* of the total sample and computed the linear regression of both distributions. For the computation of reference target salience (with one algorithm) the known luminance settings were used, for the computation of test target salience (with the combined algorithm) the known luminance settings and the matching results. The slope of the linear regression *(k_2/k_1)* was then used to weight both algorithms in the computation of predictions. Both procedures have serious disadvantages. Because they equalize maximal or mean salience values, they may mask other, perhaps important nonlinear variations in the data.

Both procedures were tested on the combinations of DARK minimum and maximum targets (Experiments 9, 12 and 13) to see if the bad predictions from *mixed combinations* can be improved. Many tested combinations include the critical algorithms 1 and 8 (see Tables 3 and 3a) and the resulting weighting factors k_1/k_2 in these combinations varied between 0.0002 (8:1) and 6100 (A:8). But the effects on MSD values were small and negligible. Neither did the best predictions notably improve nor was the ranking substantially changed. The sequence of best combinations in dense patterns, for example, was changed from 10:A, 5:A, A:1 with MSD values between 38.1 and 69.4 (Table 3a, A1a) to 10:A, A:1, 5:A with MSD values from 27.7 to 82.1 for the maximum-based weighting (i) and to A:1, 5:A, C:1 with MSD values from 40.5 to 213.3 for the linear-regression-based weighting (ii). The presumably ideal combination 10:1 of best algorithms for the two target types from section A (Table 2) produced generally even larger MSD values than without weighting factors. Because of the small effect and the fact that both weighting procedures are arbitrary and may hide interesting correlations in the data, the computation of *mixed combinations* was not changed.

With *BRIGHT targets* (Table 2, E5-H5), the best predictions from mixed combinations turned out to be better (cf. Table 3, C/D1). The best fit to the experimental data (MSD value of 5.0) was obtained for the combination of algorithm 5 (for the reference minimum target) and algorithm 10 (for the test maximum target), which both are also listed among the best performing algorithms for equal-salience matches of either target alone. But note that only algorithm 5 (for BRIGHT minimum targets) had reliably predicted the performance in equal target type comparisons (MSD value 2.2; Table 2, G5), while algorithm 10 (for BRIGHT maximum targets) was relatively poor in direct comparisons (MSD value of 44.7; Table 2, E5). A better fit should be expected if algorithm 5 were applied to both targets. This was done when testing "unique algorithms" (cf. Table 3, C2), with a relatively poor result (MSD value 77.4).

Unique Algorithms

Because of the unexpectedly poor results with mixed combinations, predictions were also made using the same, *unique algorithms* for both reference and test target salience computations (Table 3, even column numbers). In this case, the proportionality factor in the formulas is canceled. This generated a few better fits but did not generally improve the MSD values in Table 3. While for DARK target matches in Tables 3 and 3a (row A), the best predictions from unique algorithms are always better than the best predictions from mixed combinations, the opposite is true in most predictions from nonlinear luminance scales (row B) and in many predictions of BRIGHT target matches (rows C and D). Note that some unique algorithms should also have been included in the list of tested mixed combinations, like the just mentioned combination 5:5 in the matches of BRIGHT minimum vs. maximum targets. For the clarity of presentation, however, these combinations of identical algorithms are only listed as unique algorithms.

Predictions of the experimental data

We will first look at the matches of similar targets in different configurations, i.e. DARK (or BRIGHT) minimum vs. maximum targets.

Matches across minimum and maximum target configurations

The best predictions of these matches are shown in Figures 45 and 46. Remember that the experimental matches of *DARK minimum targets* in wide configurations were strongly influenced by *item salience* effects and the experimental data were split in three different sets, *dense, wide sample1* and *wide sample2* (Fig. 23). The matches in the *wide sample1* looked similar to the matches in *dense* configurations. We should expect that the different data samples are best predicted by different algorithms, as was indeed the case (Table 3a).

This is obvious in Figure 45 at labels 1 and 2. The data of target matches in dense configurations (Fig. 45a) are best, though not perfectly, fitted by predictions from algorithm 9, which however fall far off the *wide sample2* matching data in Figure 45b (label 2). (Because of their similarity with the data from dense configurations, *wide sample1* data are not plotted in Figure 45.) *Vice versa*, the best (but not perfect) predictions of matching performance in the *wide sample2* data were from algorithm 8a, which in turn was poor in predicting the performance in *dense* configurations. Predictions are, of course, identical in these two cases, as they only depend on target, distractor, and background luminance settings but not on their configuration. This difference does not show up in Table 3, where MSD computations were based on the full data set obtained with dense and wide configurations but can be seen in Table 3a, where data are split.

A similar difference is found with *mixed algorithms.* The best algorithms for target matches in *dense* configurations do not fit the matches in the *wide sample2* data, and *vice versa* (Table 3a). The individually best predictions are obtained from mixed combinations 10:A for *dense* configurations and from C:1 for the *wide sample2* data; both curves are plotted in Figure 45e (together with the matching data from dense configurations). But even these best fits from mixed combinations are worse than the according best fits from unique algorithms. The reason is seen when comparing the fits of individual curve sections (which represent matches of similar targets among different reference distractors; cf. Fig. 23). For some combinations, fits vary strongly between these sections. Predictions from the combination A:1 (red), for example, are sometimes too high and sometimes too small, and even the computationally best prediction from the combination 5:A (yellow) does fit only some sections of the data (around label 5). These variations with different curve sections are less pronounced in the predictions from best unique algorithms (Fig. 45a). In addition, the plotted predictions from mixed combinations tend to over-enhance target contrast, which is not seen in the data and not in the best fits from unique algorithms (except algorithm 6). The general over-enhancement of target contrast in mixed combinations is sometimes reduced when predictions are made on power| 0.33 luminance scales (Fig. 45f). Here, some algorithms ([8:4] and, partly hidden behind, [8:2]) produce curves that sometimes better reflect the experimental variations; this has strongly reduced the resulting MSD values (Table 3a, B1a). But even these predictions still over-enhance the differences between curve sections, and among the best algorithms listed in Table 3, B1, there are still two ([8:1], [10:A]) that over-enhance target contrast (label 6). Note that the predictions in Figure 45 are the best ones obtained, with the smallest MSD values. While it is possible that other, not tested combinations may produce better fits of the experimental data, the important point here is that mixed combinations of algorithms that are perfect or almost perfect in predicting the matching

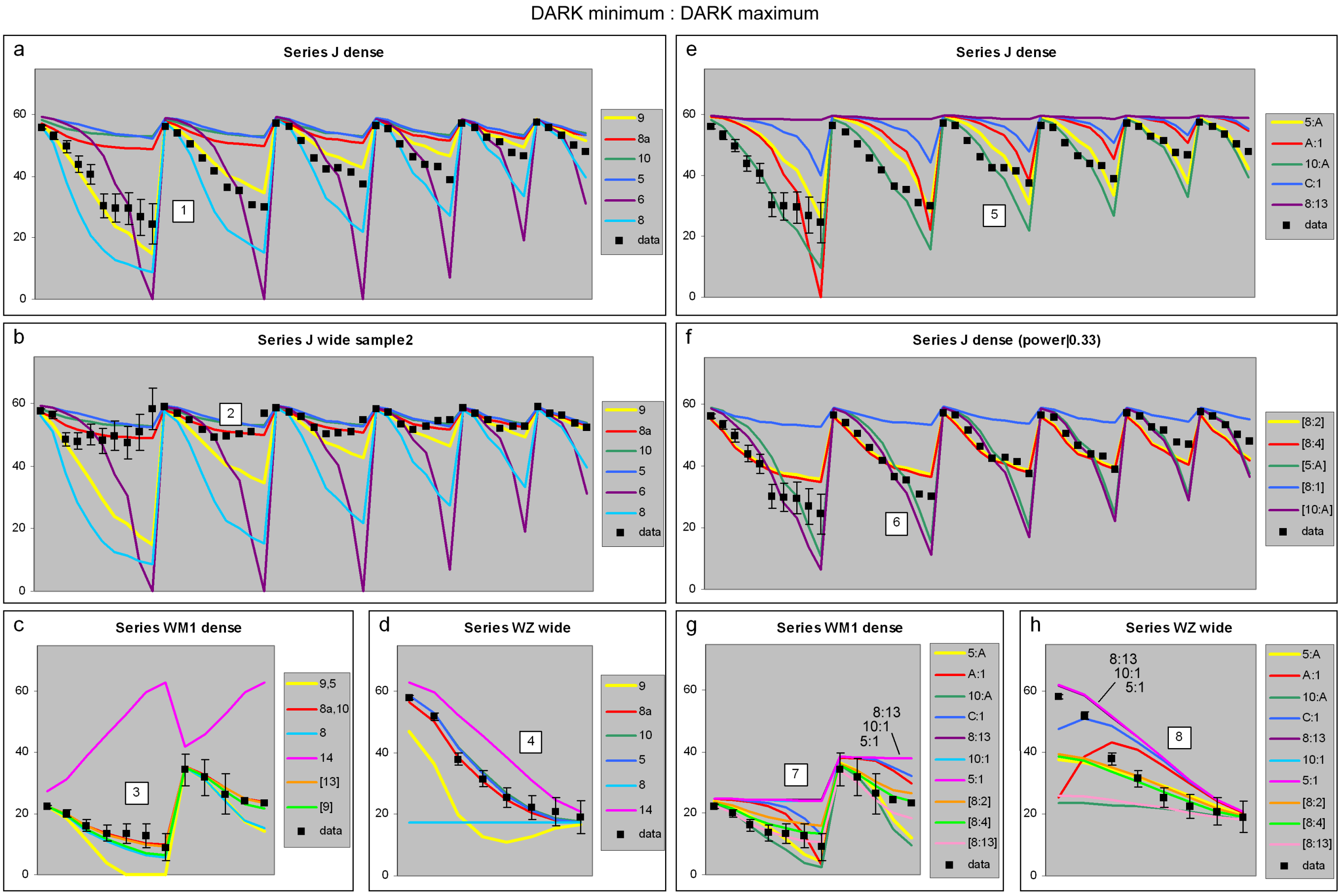


Figure 45. *Predictions of DARK minimum to maximum target matches.* Plots show matching data from section B (black squares) together with best predictions from different algorithms (colored curves). **a.-d.** Predictions from "unique" algorithms; **e.-g.** predictions from "mixed combinations"; for details see text. Many algorithms overemphasize the data variations in test series *J*. Note that data from matches in *wide* blob configurations were split (Fig.23); only the *dense* and *wide-sample2* data are here shown. Numbered labels are referred to in text. For algorithms see Table 1; brackets […] indicate that computations were performed on the power-transformed (exponent $x = 0.33$) luminance scale. Curves listed but not visible in a plot are hidden behind another, similar curve. In every graph, curves from the best predicitons (smallest MSD values in Tables 3 and 3a) are plotted yellow.

performance with the individual targets do not necessarily predict the performance in matching these targets against each other.

The curves from test series WM1 (Fig. 45c, g) and WZ (Fig. 45d, h) confirm these observations. Even the best predictions from mixed combinations do not reflect the course of data curves (labels 7 and 8), quite in contrast to the predictions from unique algorithms (labels 3 and 4). Here, like with test series *J (wide sample2)*, algorithm 8a produces the best fit to the data of test series WZ (also obtained with *wide* configurations; label 4) but good predictions are also obtained from algorithm 5 and, hidden behind, algorithm 10. In the limited luminance variations of test series WM1 (Fig. 45c), several unique algorithms produce similar predictions (cf. Table 3). The best fit is again obtained from algorithm 8a (Fig. 45c; the curve is hidden behind the predictions from algorithm [13]), whereas algorithm 9, which made the best predictions of matches in *dense* configurations and of the *wide sample1* data from test series *J* (Fig. 45a, label 1), performed rather poorly in both test series WM1 and WZ (labels 3 and 4).

Low item salience effects were less prominent in the matches of *BRIGHT minimum vs. maximum targets* (Fig. 26), and we do not have to distinguish between dense and wide blob configurations in the prediction analysis. Figure 46 shows the detailed predictions for all algorithms listed in Table 3, including algorithms applied to the power|0.33 luminance scale. The best predictions in each graph are plotted yellow, for comparison. Although quite a few MSD values from the test series *J* are small enough to indicate good fits (Table 3, C1-D2; cf. selected curves at labels 2, 7, and 8), only one *unique algorithm* makes correct predictions of all data points (Fig. 46d, algorithm [2]; cf. Table 3, D2). A few other algorithms produce nearly as small MSD values but cannot correctly predict *all* data points. For example, predictions from algorithm 10 (Fig. 46a) are often close to the data but miss the data points in the left-hand curve section near label 1. In a similar way, best predictions from *mixed combinations* (yellow curves in Fig. 46g and j) fall close to many data points (label 2) but deviate from others in a systematic way (label 8). Overall, the qualities of various fits to test series *J* reflect the same ranking as the MSD values in Table 3, C1-D2.

Also for test series *WZ*, the best fits are obtained with power|0.33 algorithms. An almost perfect fit of matches in *dense* blob configurations is produced from algorithm [2] (Fig. 46f, label 10), and of matches in wide configurations by the mixed combination [8:13] (Fig. 46l, label 12). But note that all other algorithms with best fits in test series *J* (10 in Fig. 46a; 5:10 in Fig. 46g; and [5:10] in Fig. 46j) make rather bad predictions of the matching data in test series *WZ* (Fig. 46c, i, and l, labels 4, 6, and 13).

In test series *WM4* predictions from several algorithms again collapse (cf. Table 3, C/D4) and only two valid MSD computations from unique algorithms could be made (Fig. 46b). One algorithm (1=2=5=9; yellow) produced a particular good fit of the data, the other algorithm (11=12=14; blue) no fit at all (label 3). The graph shows a third curve from algorithm 10, also with a rather bad fit, which is only plotted for comparison with Figure 46a, where predictions from algorithm 10 had produced a particularly good fit of the data. For MSD computation in test series WM4 (Fig. 46b), this algorithm was classified invalid because one of the nine data points fell outside the accepted luminance range (1/9 > 10%). Also with mixed combinations, good fits with test series *J* (e.g., C:5, label 2) do not necessarily predict good fits with other test series (e.g., Fig. 46h, C:5, label 5). Best predictions from power| 0.33 transforms (Fig. 46e, k; labels 9 and 11) were generally worse with this test series.

Figures 45 and 46 illustrate three important aspects of these predictions, (i) their variability across different test series, (ii) the, in principle, similar efficiency of unique algorithms and mixed combinations, and (iii) the occasional but not general improvement of fits when replacing the linear by a power|0.33-transformed luminance scale. Although the best fit of BRIGHT minimum and maximum target matches in test series *J* was obtained from a (unique) power|0.33 algorithm (Fig. 46d), quite good and in some test series even better predictions were obtained from algorithms on the linear luminance scale or from mixed combinations. As a whole, the MSD values (Table 3) are similar for unique and mixed combinations and for linear and nonlinear luminance scales. Thus there is no general advantage of using power-transformed instead of linear luminance scales or of using unique instead of mixed algorithms when predicting target matches of similar luminance polarity.

Matches across different luminance polarities

This pattern changes with cross-polarity matches (Table 4). Here, best predictions from unique algorithms on linear luminance scales are obtained with algorithm 3,

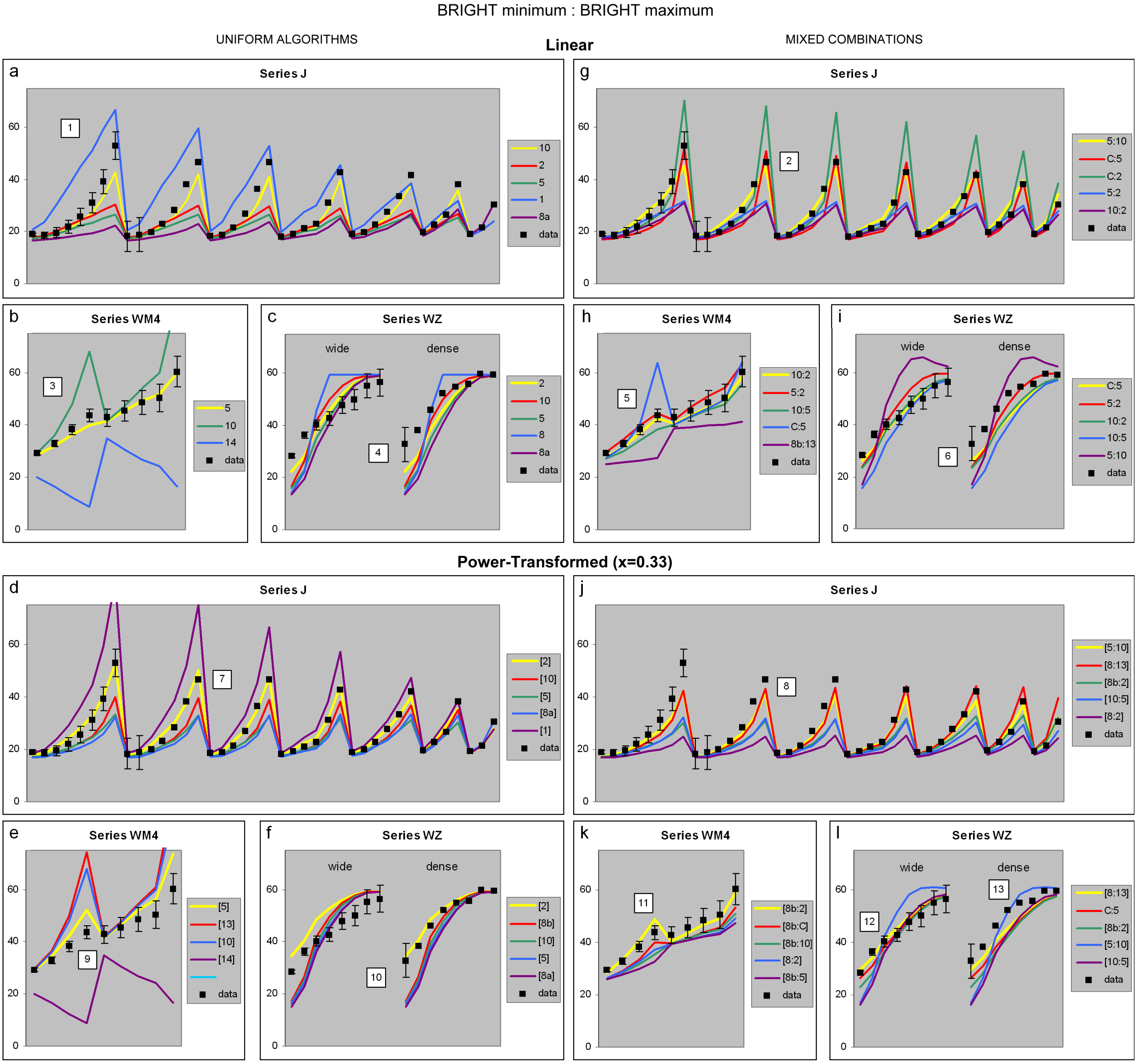


Figure 46. *Predictions of BRIGHT minimum to maximum target matches.* Plots show matching data from section B (black squares) together with best predictions from different algorithms (colored curves). **a-f.** Predictions from "unique" algorithms on linear *(a-c)* and power|0.33-transformed scales *(d-e)*; **g-l.** according predictions from "mixed combinations" (*g-i,* linear; *j-l,* power|0.33 transforms); for details see text. Good fits in certain test series are often singular results that do not transfer to other test series. Best predictions in each graph are plotted yellow, for easier comaprison, and second-best ones red.

which produces smaller MSD values than the other unique algorithms (Table 4, A2 and A4). Most other "best" predictions from the linear luminance scale are worse than the predictions from the power|0.33 luminance scale (Table 4, B2 and B4). This is particularly obvious with DARK to BRIGHT maximum target matches (Table 4, A/B2) where the difference holds over all listed values and where several algorithms on the power-transformed scale made even better predictions than algorithm 3 (which is “immune” against power transformations).

The differences are obvious when looking at the quality of predictions in detail (Fig. 47 and 48). With test series block *WB*, for example, only the predictions from algorithm 3 fall sometimes close to the data (Fig. 47a) even though the contrast of very bright (Fig. 47a, label 1) or very dark targets (Fig. 48a, label 1) is systematically exceeded. Linear predictions from the other algorithms show larger and more frequent deviations from the data. This changes when computations are based on a power-transformed luminance scale (Fig. 47c, label 4, and 48c, label 3). Now quite a few other predictions also fall close to the data points, some even closer than the predictions from algorithm 3. The difference is less obvious but also present in the predictions from mixed combinations (Fig. 47b, d and 48b, d). Here, most predicted curves partially deviate from the data (Fig. 47, labels 2 and 3; Fig. 48, label 2), but predictions based on the power|0.33 transform generally fall closer to the data points (Figs. 47d, label 5, and 48d, labels 4 and 5). The advantage of power|0.33-transformed over linear luminance scales in mixed combinations is also seen in Table 4; all MSD values from nonlinear luminance scales are smaller than the values from linear scales at corresponding ranking positions (cf. Table 4, B1 vs. A1 and B3 vs. A3).

Remember that the *test conditions* (not the data) in series *WB2* were, in principle, mirrored from the test conditions in series *WB1*. What should that imply for the analysis of best fitting algorithms? First of all, if the selection of mixed combinations would at all be meaningful, one should expect that the best combinations in one test would simply be reversed in the other test. This was not the case, neither for the best algorithms on the linear nor for those on the nonlinear luminance scale (cf. Table 4). None of the best combinations listed in one test showed up, in reversed form, in the list of best combinations for the other test. The fact that some algorithms were listed in mirrored positions but in

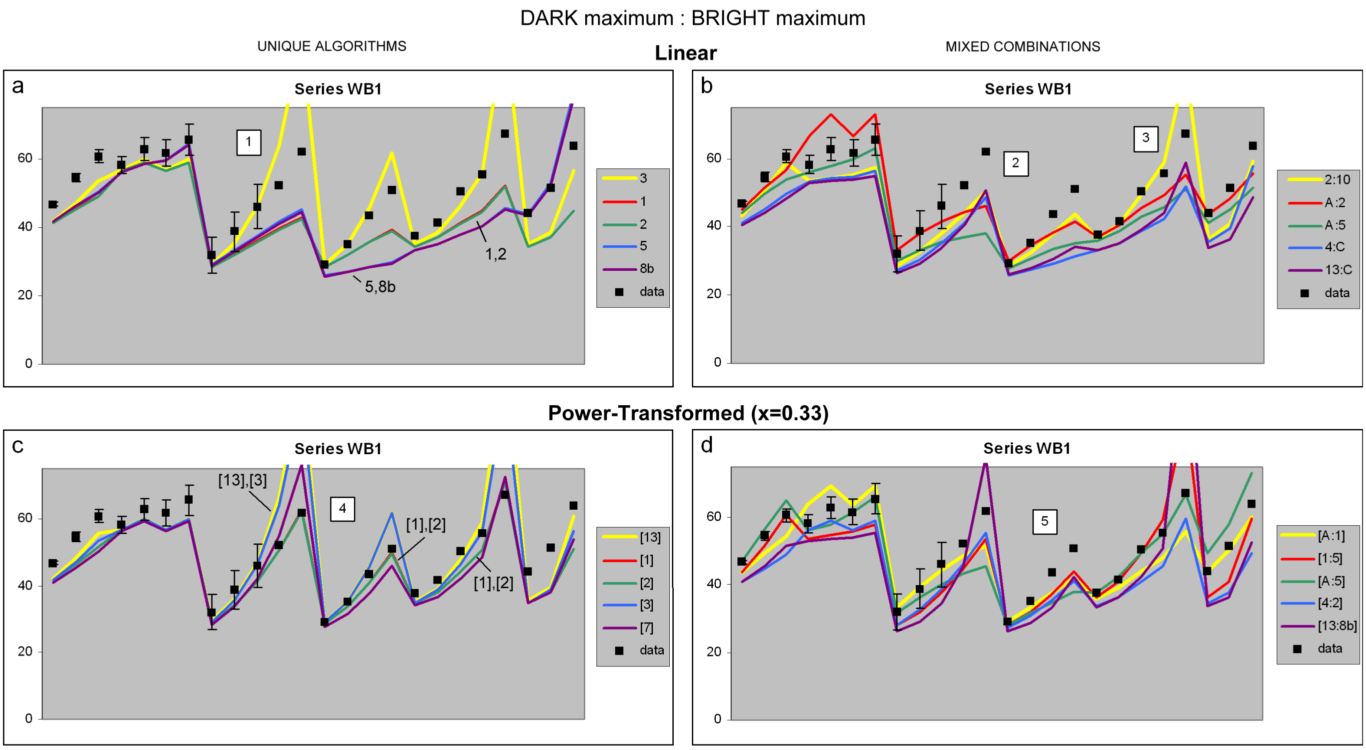


Figure 47. *Predictions of DARK (reference) to BRIGHT (test) maximum target matches in Experiment 11 (test series WB1).* Data (black squares) and predictions from all algorithms in Table 4 (colored curves); presentation as in the previous figures. For details see text.

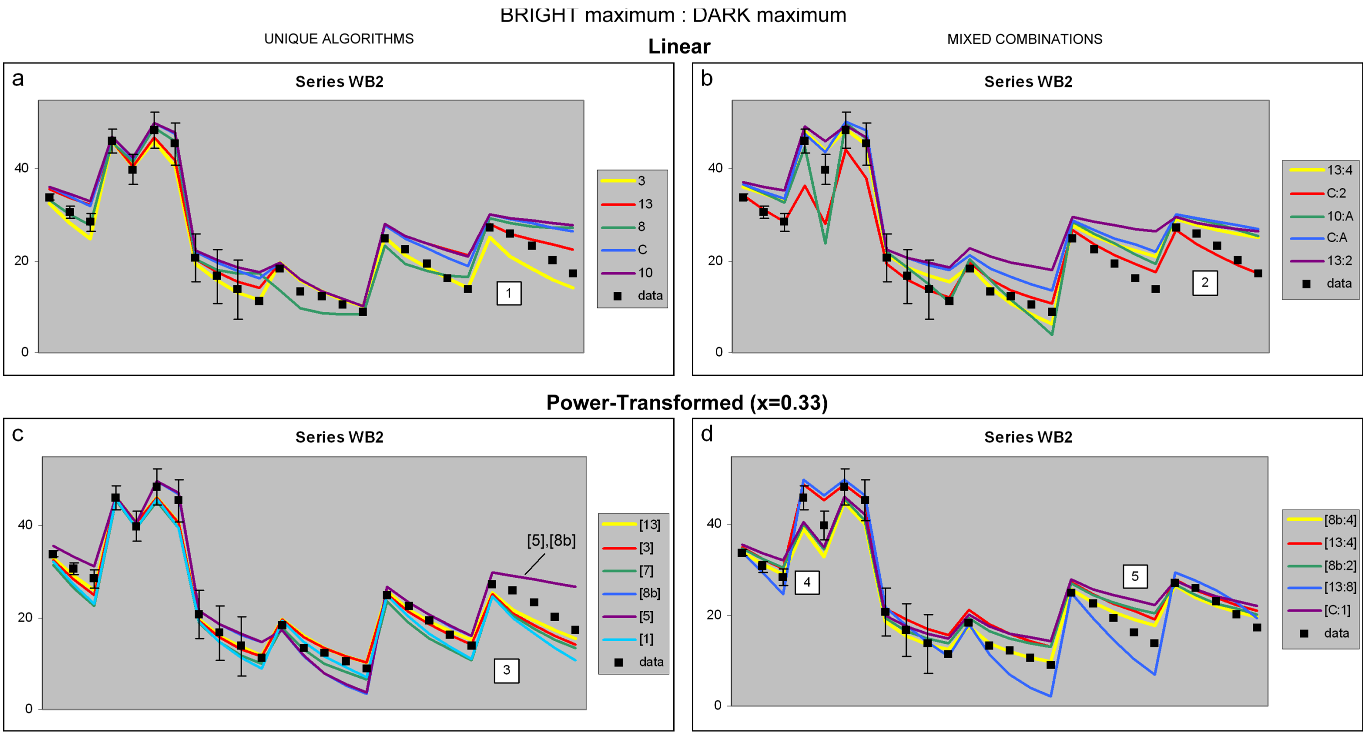


Figure 48. *Predictions of BRIGHT (reference) to DARK (test) maximum target matches in Experiment 11 (test series WB2).* Data (black squares) and predictions from all algorithms in Table 4 (colored curves); presentation as in the previous figures (but note the different scales). For details see text.

combination with other algorithms (e.g. 2:10 and 10:A) is simply due to the fact that exactly these algorithms had been selected for the analysis of mixed combinations; the same lists of Table 2 were used for combinations of identical target types. The failure to find best fits from reversed combinations in test series *WB1* and *WB2* is surprising and suggests that the exact selection of algorithms in mixed combinations is far less important for the quality of fits than one might expect. A second expectation from reversed test conditions in test series *WB1* and *WB2* should be the consistence of best fitting unique algorithms; an algorithm that could reliably predict the performance of target matches in one test series should also reliably predict the performance in reversed matches. For unique algorithms, there is indeed a high consistency between best algorithms in mirrored test series, but only for power|0.33 algorithms. Four of the five best algorithms occur in both ranking lists (Table 4, B2 and B4). With the linear luminance scales (A2 and 4) there is - beside algorithm 3 (Michelson contrast) - no overlap.

MSD distributions

On a first view one might have assumed that *mixed combinations* of algorithms should make better predictions of the experimental matches than *unique algorithms*, in particular when algorithms are combined that closely predict the salience properties of the involved targets. But, as we have seen, the computational results do not support this assumption. Although predictions from mixed combinations were quite good in some cases, in other cases the predictions from unique algorithms might have been better. In general, neither the MSD values in Table 3 nor the detailed inspection of predictions in Figures 46-48 have revealed any systematic advantage of one type of computation over the other.

This is also seen in the MSD distributions of Figure 49. When ordered for magnitude, MSD values from unique algorithms (lightly hatched bars) and mixed combinations (densely hatched bars) intermingle; there is no indication that mixed combinations (even not of algorithms optimal for the compared targets) make generally better predictions of the data than unique algorithms. In the matches of DARK and BRIGHT maximum targets (Fig. 49b), the best predictions were obtained from either a mixed combination (upper graph) or a unique algorithm (lower graph), but the best predictions from the other type of computations were always close to these values. The same is true for the predictions of minimum-to-maximum target matches in Figure 49a.

Note that the argument is based on the intermingled *sequence* of MSD values when sorted for magnitude; the sometimes unequal counts in the distributions must not confuse. Only in some graphs of Figure 49 were about as many values from unique algorithms as from mixed combinations; in other distributions (e.g., the lower graph of Fig. 49b) values from mixed combinations predominate. This is due to the different numbers of tested algorithms. While Table 1 lists 24 unique algorithms, some of which are mathematically identical in certain tests and then shown only once in the distributions, the number of tested mixed combinations can be larger and does not include mathematically identical computations. Both types of computations generated quite a few predictions outside the experimental luminance range, so that the according MSD values had to be set invalid. In addition, the graphs in Figure 49 show only the best MSD values; particularly large values are not plotted.

Really good predictions (MSD values < 10) are only seen with the BRIGHT to DARK maximum target matches (Fig. 49b); in all other matches the best

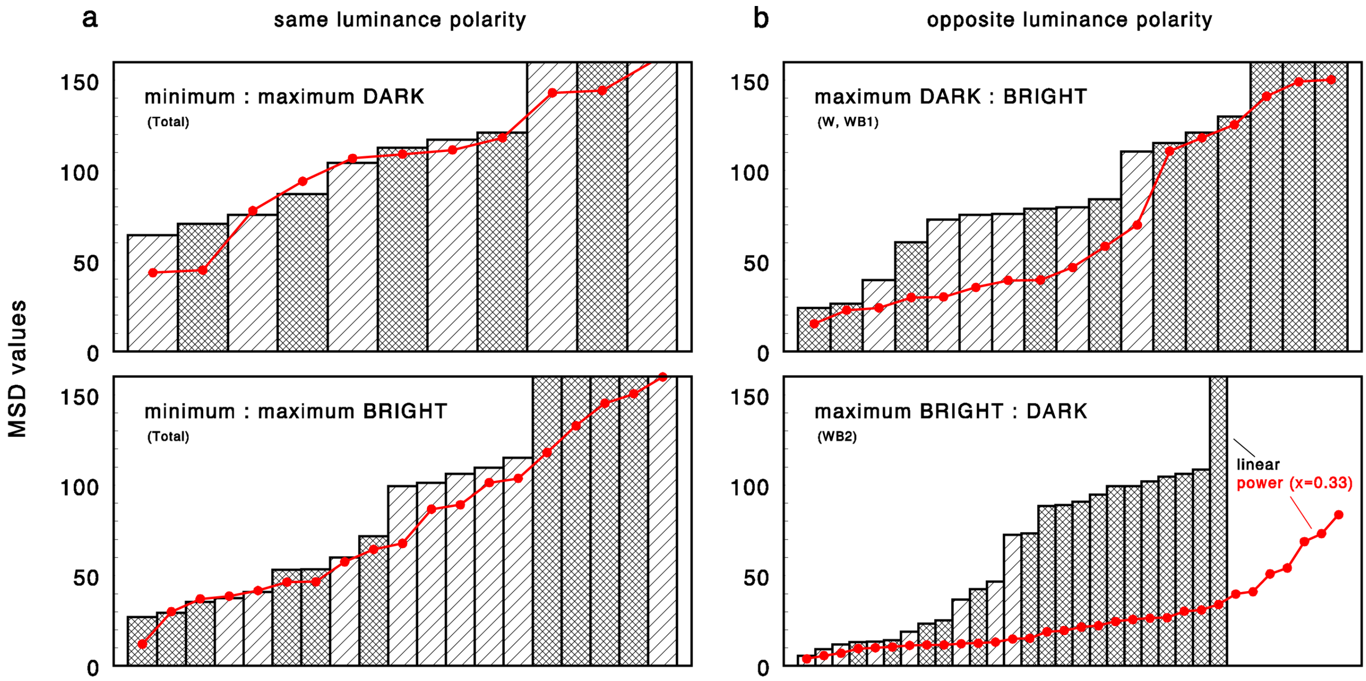


Figure 49. *Distributions of MSD values in section B experiments.* **a.** MSD values for DARK (top) and BRIGHT (bottom) minimum to maximum target matches accumulated from Experiments 9-10, 12 and 13. **b.** MSD values for DARK to BRIGHT (top) and BRIGHT to DARK target matches (bottom) in Experiment 11. Histograms give the MSD values from linear analysis for *unique algorithms* (lightly hatched) and *mixed combinations* (cross-hatched). Red curves show the MSD values from power|0.33 analysis (not distinguished for unique algorithms and mixed combinations). Both distributions, from linear and power|0.33 MSD analysis, are independently sorted for increasing values; same positions thus do not necessarily refer to the same algorithm (different to the presentation in Fig.43). In all graphs values are plotted until both distributions exceed a value of 150; larger MSD values are not shown. In the bottom graph of *(b)* there were many more valid MSD values in the power|0.33 analysis (red) than in the linear analysis (histogram); these values are added on the right-hand side. The graphs illustrate two major results from section C. Predictions from unique algorithms (lightly hatched) and mixed combinations (cross-hatched) are not systematically different. Predictions from linear and power|0.33 analyses differ only in the cross-polarity matches of DARK and BRIGHT targets *(b);* predictions from the power-transforms then produced, on average, smaller MSD values. When targets were similar and both either DARK or BRIGHT, distributions of MSD values were more similar *(a)*.

predictions are notably worse. For the DARK minimum to maximum target matches (upper graph in Fig. 49a) this is partly due to the averaging of data from dense and wide configurations, which had produced different matches. The much smaller MSD values of the individual fits (Table 3a) are here not shown.

Predictions from nonlinear luminance scales

The red curves in the graphs of Figure 49 show the MSD distributions from nonlinear analysis, when the same algorithms were tested after a power|0.33 transformation. In both analyses, there were invalid data so that the two MSD values from the same algorithms could often not be compared. Therefore, different to Figure 43, MSD distributions from linear and nonlinear luminance scales are here plotted independent of each other. Both show the same count of best MSD values ordered for magnitude, but data points on the red curves and the according entries in the bar histograms may come from different algorithms.

The presentation illustrates an important point. While the MSD distributions from linear and power|0.33 computations are very similar in the matches of similar targets in different configurations (Fig. 49a), the distributions strongly differ in the matches of targets with different luminance polarity (Fig. 49b). In the latter matches, there are many more small MSD values from power|0.33 algorithms (red curves) than from power|1.0 algorithms (bar histograms). The differences between the distributions are visualized in Figure 44.

Thus, different to the little importance of nonlinear luminance scales in the previous analyses, predictions of cross-polarity matches are generally improved by power|0.33 transforms. Nonlinear luminance scales are important when DARK and BRIGHT targets are compared for salience (“cross polarity matches”) but not or less so when similar targets within the same luminance polarity are compared, i.e. when both targets are either DARK or BRIGHT. An interesting behavior in this aspect shows algorithm 3 (Michelson contrast). It does already implement a magnitude-dependent scaling of luminance and is not affected by power transforms of the luminance scale. Predictions from algorithm 3 were among the best ones in test series block *WB* (cf. Table 4) but were still topped by some power|0.33 algorithms.

In summary, we have seen that matches of different target types were sometimes difficult to predict. Combinations of algorithms that were optimal in predicting the salience characteristics of either target alone, often failed to predict the matching performance with target combinations. In most cases, these *mixed combinations* were not better than *unique* algorithms in which salience matches were predicted from one common algorithm, irrespective of whether that could adequately predict the salience characteristics of either target alone. This raises questions about the physical basis and plausibility of fitting algorithms. Another important finding is that the prediction of salience matches between BRIGHT and DARK targets, i.e. between targets at different luminance polarities, was strongly improved by the usage of power-transforms. Best results were obtained with a power|0.33 transform. This is in strong contrast to matches of similar targets, which could reliably be predicted also from linear analysis.

Conclusions from computations part II:

Best predictions of equal-salience matches of different targets are often singular and cannot be transferred to other stimulus conditions. – Mixed combinations of algorithms that reliably predict the salience of the two targets are not better in predicting combined target matches than other, arbitrary algorithms. – Predictions of cross-polarity target matches are generally improved by the use of power functions and follow Stevens’ brightness law.

Discussion of section C

Hints for reading: Congratulations; you have passed the hardest part. The major findings of section C will now be summarized: (i) there is no common super algorithm but different target combinations are predicted by different algorithms; (ii) algorithms on power-transformed luminance scales help to predict cross-polarity (dark vs. bright) but not within-polarity matches (dark vs. dark or bright vs. bright). After discussion of the general properties of the tested algorithms, the best algorithms for each target type are summarized.

Although not all matches of sections A and B were predicted exactly, the analysis has nevertheless revealed interesting results. First, computations did not identify a single “super” algorithm that could predict equal-salience matches in all test conditions. Instead, the quality of predictions from the same algorithms differed between target types and target configurations. Second, in spite of this variability, computations helped to rate algorithms that did predict, and others that failed to predict, the equal-salience matches in a particular test condition. Third, computations did, in particular, identify test conditions in which matches were better explained by a power-transformed than the linear luminance scale.

No "super-algorithm" for luminance-defined salience

One reason to start the thorough analysis of various computational algorithms was the hope to find an algorithm that could account for *all* salience variations measured in sections A and B. Contrary to this hope, however, the quality of predictions from most algorithms varied considerably when applied to different targets or targets in different configurations. Algorithm 1, for example, that had closely fitted the equal-salience matches of DARK maximum targets, produced rather poor predictions of BRIGHT maximum target matches (unless targets were the brightest items in the scene). Algorithm 5, on the other hand, which gave the best fits to matches of BRIGHT targets in either maximum or minimum configuration, failed to predict the matches of targets in exactly this combination. Thus, the computations in section C contradict the idea of a super-algorithm for luminance-defined salience and instead show that there are many different algorithms that may fit the data of a certain equal-salience match and fail in others. This is particularly obvious in the good performance of hybrid algorithms tested on BRIGHT maximum target matches (Table 2H). The simple averaging of two computations fitted the experimental data of some test series better than did either of the tested algorithms alone.

It is unlike however that the super-algorithm exists and was not found because it was just not included in the list of Table 1. All standard contrast algorithms were tested; differences, Weber contrast, Whittle contrast, Michelson contrast, and various ways of normalizing the target-distractor difference. As a whole, algorithms covered a wide spectrum of target, distractor and background combinations. Rather, experiments have provided strong evidence that no such super-algorithm for luminance-defined salience exists, at least none that would only be based on the luminance settings of targets, distractors, and background. Many experimental matches were, in fact, closely predicted by certain algorithms, which however failed when the same target was presented in a different configuration or was matched with another target.

General properties of tested algorithms

Given this variability, it might be interesting to evaluate the general performance of algorithms and look at principle differences in their predictions. Note that the algorithms of Table 1 can be grouped by the pattern components they take care of. One major group (algorithms 1 to 3) represents pure measures of target-to-distractor differences (discrimination salience) without implementing background luminance as a parameter. Some of these algorithms produced rather good fits of the data. In certain matches, however, the perceived target salience *did* depend on background luminance and algorithms that had implemented this parameter made better predictions of the experimental data. Other matches, in wide raster configurations, also depended on item salience, and item-to-background differences should be included to correctly predict *these* data. And target salience might even depend on the comparison of various salience components, which can only be predicted if they are all implemented in the presumed salience computation. Can we draw *general* conclusions about predictions from these different algorithms?

Role of distractors. Although observers only had to match *targets*, distractors clearly affected their performance. Two algorithms that exclusively encode target-to-background contrast but not distractor settings (algorithms 14 and 15) generally failed to predict the obtained matches.

Role of background. Background settings, on the other hand, were sometimes less important. Note that algorithms 1-3 which do not care of background luminance are frequently listed among the best algorithms in Tables 2-4. This is not true for minimum targets, however, the salience matches of which usually depended on both background and distractor settings. And it is not strictly true for BRIGHT targets, the salience matches of which were often seen to vary with background luminance, although predictions from algorithm 2 were often quite good (e.g., Fig. 3).

Comparisons of item salience. If we assume that the *item salience* of targets and distractors is encoded by their Weber (or Michelson) contrast to background, as was generally found for single items or pure distractor arrays (Nothdurft, 2015), one might propose two principle ways of how to compare these values. The item salience measures could be *subtracted*, as in algorithms 4 and 13 (for Weber and Michelson contrasts, respectively), or be *divided,* as in algorithms 11 and 12. Algorithm 11 is mathematically equivalent with the *salmin* algorithm 5, which is frequently listed in Table 2. But for the Michelson contrast-based computations, differences (algorithm 13) are far more frequent than ratios (algorithm 12).

Comparisons of discrimination and item salience. Another interesting combination of target properties to look at is the comparison of target-to-distractor and target-to-background differences. Here, the *ratio* of values (algorithms A and C) had provided reliable predictions, but not their *differences* (algorithms B and D). That is, item salience and discrimination salience of the target are apparently not subtracted to calculate its apparent salience. But good fits from algorithms A and C are also not very frequent and one cannot conclude that a target's salience is generally computed as the ratio of discrimination and item salience. However, the ratio of target-to-distractor and *distractor*-to-background differences (algorithms 9 and 10) produced good singular fits (particularly in the matches of DARK minimum and maximum targets).

Normalization. An important aspect of salience estimates is the normalization of luminance differences. Algorithms 1-7 (including Weber, Whittle, and Michelson contrasts) differ only in their normalization (which is even absent in algorithm 1). In algorithms 8, 8a, and 8b, the Michelson or Weber contrasts themselves are normalized to either the entire luminance span of the pattern (algorithms 8 and 8b) or to the predominant luminance span given by background and distractor luminance (algorithm 8a). Normalizing the Michelson contrast has been reported to be critical for correct predictions of transparency (Singh & Anderson, 2006) and did also reliably predict the matching performance in several target combinations in the present study.

Equal-salience matches do not measure the strength of salience. The good performance of algorithms A and C with minimum targets underlines a principle limitation of the computational analysis. These two algorithms compute salience as the ratio of target-to-distractor (discrimination salience) and target-to-background contrasts (item salience). Since minimum targets vary between background and distractor luminance, either one of the two differences may become very small, and hence their ratio very large. It is unlikely, however, (and was never observed in the experiments) that a minimum target had become extremely salient when approaching background luminance. The opposite was found (and discussed above); targets very similar to the background, in fact, *loose* their item salience and are hard to be seen, quite different to the salience modulation predicted from algorithms A and C. Why did these algorithms then produce so good fits? The fits were good only in the equal-salience matches of section A (Table 2, rows C and D) where similar test configurations were compared without measuring the strength of each target's salience. When the measure of salience itself became important, as in the matches of section B, predictions from these algorithms were poor and did not reach the best-of lists of unique algorithms in Table 3, rows A and B. This indicates the general limitation of equal-salience matches in section A. Good predictions in terms of small deviations of predicted from experimental data *per se* do not identify the underlying mechanism. The MSD can only measure the quality of a particular algorithm to predict the experimental data but cannot identify the validity of this computation. If stimulus parameters between reference and test patterns change in proportion, as in the experiments of section A, the fit may be perfect even if the computed salience is quite different to what is perceived. To identify the algorithms that compute salience, we thus must also look into their performance in detail and verify that they indeed behave the same way as perception does.

Best salience algorithms

I shall now summarize the findings of section C for different target types. But before going into details of the computations, it seems helpful to recall some important properties of the tests and computations.

1. DARK and BRIGHT targets do not necessarily display different luminance but are named so because of their different luminance *polarity to background*. In comparison to distractors, however, minimum targets have a different polarity than maximum targets, DARK minimum targets are *brighter*, and BRIGHT minimum targets *darker*, than distractors.

2. Weber contrast (algorithm 2) and Michelson contrast (algorithm 3) make the same predictions for equal-salient targets that have the same luminance polarity to the

comparator. For item salience (target-to-background differences) that would be the case when both targets are DARK or BRIGHT but not when DARK and BRIGHT targets are matched. For discrimination salience (target-to-distractor differences), however, the two algorithms make identical predictions only when targets have the same luminance polarity to distractors, e.g. when maximum targets are matched to maximum targets, or minimum targets to minimum targets. If this is not the case, predictions from the Weber contrast and the Michelson contrast will differ. In all other conditions, however, a good fit of predictions from the Weber contrast does not indicate that it would be better suited than the Michelson contrast, and *vice versa*.

This is quickly verified by computing the luminance of an equal-salient target, tg_2, from the according algorithms in Table 1. For the combination of two BRIGHT maximum targets (similarly two DARK maximum targets) we obtain

$$\text{(Weber)} \qquad \frac{tg_1 - dis_1}{dis_1} = \frac{tg_2 - dis_2}{dis_2} \quad \Rightarrow \quad tg_2 = \frac{dis_2}{dis_1} \cdot tg_1$$

$$\text{(Michelson)} \qquad \frac{tg_1 - dis_1}{tg_1 + dis_1} = \frac{tg_2 - dis_2}{tg_2 + dis_2} \quad \Rightarrow \quad tg_2 = \frac{dis_2}{dis_1} \cdot tg_1 \ ,$$

but for the combination of a BRIGHT and DARK maximum target

$$\text{(Weber)} \qquad \frac{tg_1 - dis_1}{dis_1} = \frac{dis_2 - tg_2}{dis_2} \quad \Rightarrow \quad tg_2 = 2 \cdot dis_2 - \frac{dis_2}{dis_1} \cdot tg_1$$

$$\text{(Michelson)} \qquad \frac{tg_1 - dis_1}{tg_1 + dis_1} = \frac{dis_2 - tg_2}{dis_2 + tg_2} \quad \Rightarrow \quad tg_2 = \frac{dis_1 \cdot dis_2}{tg_1} \ .$$

Thus, while for computations based on the Michelson contrast the ratio of the brighter to the dimmer item is maintained,

$$\frac{tg_1}{dis_1} = \frac{tg_2}{dis_2} \quad \text{and} \quad \frac{tg_1}{dis_1} = \frac{dis_2}{tg_2} \ \text{, respectively,}$$

for computations based on the Weber contrast the ratio of target to distractor luminance remains important, or eventually its complement to 2,

$$\frac{tg_1}{dis_1} = \frac{tg_2}{dis_2} \quad \text{and} \quad \frac{tg_1}{dis_1} = 2 - \frac{tg_2}{dis_2} \ .$$

3. The evaluation of predictions was largely based on the analysis of MSD values, which tell nothing about the underlying mechanism (as discussed above with algorithms A and C). Several examples in section C have shown that even curves with opposite modulation may generate similarly small MSD values and may thus indicate similarly good fits, which are however not based on similar mechanisms.

DARK maximum targets

The overall best predictions of matching data in section A were obtained from algorithm 1, the *constant-addition principle*. Targets appeared equally salient among distractors when their luminance difference to distractors was constant. This finding is remarkable in two aspects. First, it stresses the little importance of background luminance in the salience computation of DARK popout targets (which was tested in Experiment 4 and is demonstrated in Fig. 14). Second, the finding is contrary to what one might have expected from natural viewing conditions, where predominant luminance variations come from illumination changes and result in *constant-ratio* luminance changes of purely reflective surfaces. If item and discrimination salience had to be constant under different illuminations, one might expect that salience computation should follow algorithm 2 rather than algorithm 1. In natural scenes, constant addition (algorithm 1) should only occur with self-luminous items which typically are *brighter* than their surroundings. (It is not easy to imagine dark equivalents that would subtract a constant amount of luminance. The only examples I could think of are holes, which would however show no surface reflection and hence no additive component at all.) It should be stressed though, that the luminance variations in the present study were arbitrary and usually not consistent with a natural illumination change; background and distractor settings have been varied independently all over the available luminance ranges.

On a first glance, however, algorithm 2 (constant ratio) did apparently make quite good predictions of the data in some test series (cf. Table 2, row A). But this impression is misleading. In two series (*K* and *F)*, predictions from algorithm 2 (and 3) should be (about) as good as predictions from algorithm 1, because distractors were (about) identical (cf. *equation 2)*. In test series *L,* the better performance of algorithm 1 was simply not listed because it generated too many invalid predictions and hence was excluded from MSD analysis. Careful inspection of Figures 5 and 6, however, suggests that predictions from algorithm 1 were not always bad in these test series but mainly deviated from the data at very low luminance settings. Altogether, in the analysis of all matches with DARK maximum targets, predictions from algorithm 1 were better than predictions from algorithm 2 (Table 2, A5). Particularly in test series *L* and *J*, also predictions from algorithm 8 (Singh & Anderson, 2006) were quite good (and almost indistinguishable from the predictions from algorithm 1 in Fig. 17) but the general form of predictions (Fig. 18) does not seem to reflect the measured data (Fig. 19). In the Totals, this algorithm performed worse than algorithm 1.

Fits did generally not improve when nonlinear instead of linear luminance scales were used, except for matches of DARK and BRIGHT maximum targets (Table 4). Here, predictions based on power|0.33-transforms of luminance were generally better than predictions based on the linear luminance scale. For DARK to BRIGHT target matches, predictions from algorithms [1] were the best; for BRIGHT to DARK matches they were among the best ones listed in Table 4.

Despite the good performance of algorithm 1 in matches of DARK maximum targets, the algorithm was apparently less valid in predicting the matches of DARK minimum to maximum targets (Table 3, row A). When looking at these matches in the scatter plots of Figure 27, however, a linear but differently weighted relationship is obvious. Equal-salient luminance variations are reduced to about 60% in maximum targets.

BRIGHT maximum targets

The original suspicion (Fig. 3) that salience matches of BRIGHT (maximum and minimum) targets might follow the *salmin* algorithm 5 was confirmed. Predictions from algorithm 5 were the best, or second-best, in almost all top-five listings of Table 2 (rows E and G). Only in test series block *L* it had to be set invalid, since more than 10% of the predictions fell outside the tested luminance range. Here, best predictions were obtained from the *constant-ratio principle*, algorithm 2, which also made good predictions in most other test series. As can be seen from the formulas in Table 1, the two algorithms are similar for low background luminance (for zero backgrounds they are identical); thus their similar performance is not surprising. In spite of the good performance of these two algorithms with BRIGHT maximum and BRIGHT minimum targets (Table 2, E5 and G5), predictions from these algorithms were poorer in direct matches of these two targets (Table 3, rows C and D) except in test series WM4. Even though both algorithms are listed among the top fives with series *J*, performance was better with algorithm [2]. Slightly better predictions of BRIGHT maximum target matches in test series *K* and *F* (cf. Table 2) were obtained from algorithm 8b (a modification of the *Singh* algorithm 8 based on Weber instead of Michelson contrast). Other variants of that algorithm, algorithms 8 and 8a, also occurred in the top-five lists of Table 2.

Overall, the predictions of BRIGHT maximum target matches were poorer (larger MSD values) than the predictions of DARK maximum target matches. This was changed by the use of hybrid algorithms, in which predictions from two selected algorithms were averaged (Table 2H). Three standard algorithms were chosen for these averages, the *constant-addition principle* (algorithm 1), the *constant-ratio principle* (algorithm 2), and the *salmin* algorithm 5. Some of these averages produced notably better predictions than the algorithms themselves. This suggests that subjects might have made their adjustments along two lines of computations. In particular, combinations of algorithms 1 or 2 and algorithm 5 (<1,5> and <2,5>) made often good predictions of the data, whereas the combination of constant-addition and constant-ratio (<1,2>) did usually not improve the fits.

As a whole, predictions based on nonlinear luminance scales were not better than predictions based on the linear scale for matches of BRIGHT targets. Certain algorithms, however, made better predictions when applied to the power|0.33 scale (Table 2, E/F5). For example, algorithm 1 made poorer predictions (MSD value 71.8, not listed) than algorithm [1] (MSD value 34.9). But a strong general improvement from replacing the linear luminance scale by the power|0.33 transforms was only seen in the cross-luminance-polarity matches of DARK and BRIGHT maximum targets (Table 4).

DARK minimum targets

Predictions of minimum target matches were generally very good, despite the occurrence of local deviations as demonstrated in Experiment 5 (Fig. 10). The best algorithms for DARK minimum targets (Table 2, row C) were algorithms 10, A, C, 12, and 5 in all test series except series *F* (where the latter three algorithms in the list are replaced by algorithms 13, B, and 9), but a ranking of predictions with such small deviations from the data might not be meaningful. The quality of fits was reduced in test series block *O* but the ranking of best algorithms was still similar to that in test series block *K*.

Given the already low MSD values obtained from linear computations, it is not surprising that predictions were not improved when the power-transformed instead of the linear luminance scale was used (Table 2, row D).

From this high quality of predictions for DARK minimum target matches it is surprising that predictions of DARK minimum to maximum target matches were so much poorer (Tables 3 and 3a). The best predictions of different target matches were obtained from algorithms 9 (dense configurations) and 8a (wide configurations, sample2 data), which were both not listed as particularly successful in the best-of lists for DARK minimum or DARK maximum target matches themselves (Table 2, A/C5).

BRIGHT minimum targets

Similar observations are made with the predictions of BRIGHT minimum target matches. MSD values were generally small and similar across the different test series (Table 2, rows G and H). About the same algorithms made the best predictions on linear and power-transformed luminance scales, with only little if any systematic improvements on either scale. Different to DARK minimum target matches, however, there is more consistency between best algorithms for BRIGHT maximum or minimum target matches (Table 2, E5, G5) and best algorithms for BRIGHT minimum vs. maximum target matches (Table 3, C7-C8), even though MSD values differ. Both mixed combinations and unique algorithms in Table 3 list several algorithms that are also listed in Table 2, rows G and H.

Matches of different target types

Close predictions of the experimental data (small MSD values) were generally rare in the matches of different target types (Tables 3 and 3a). Good fits were highly selective and often found for only one or two algorithms, sometimes then even restricted to matches in dense or wide blob configurations; most other algorithms produced large deviations from the experimental data. For *DARK minimum vs. maximum target matches*, only one unique algorithm (algorithm 9) produced a reasonably small MSD value for matches in dense configurations of test series *J* (Table 3a, A2a and b); this algorithm could however not reliably fit the similar matches in other test series (cf. Fig. 45c and d). For matches in wide configurations, which strongly suffered from item salience effects, several algorithms produced reasonable MSD values (Table 3a, A2c), but these matches were only little modulated (cf. Fig. 45b) and good predictions might thus not be meaningful. Also in mixed combinations, good matches if found at all were singular events; most predictions were rather bad.

The same is true for *BRIGHT minimum to maximum target* matches. The experimental data were closely predicted by only few algorithms (e.g., unique algorithms 10 and [2]), and small MSD values are rare in the lists of Table 3. It is interesting to note that the unique algorithm 5, which made the best predictions of BRIGHT minimum and maximum target matches (Table 2, E/G5), was relatively poor (MSD values > 50) in predictions of both targets matched to each other, except in series *WM4* (Table 3, C4).

Finally, even the matches of *DARK to BRIGHT maximum targets* (and *vice versa)* were not perfectly predicted (Figs. 47 and 48; Table 4). Best predictions were obtained from algorithms on the power|0.33-transformed luminance scale. The most important observation here is that predictions based on power-transforms were generally better than predictions from the linear luminance scale.

Generalization and Conclusions

Several conclusions can now be drawn from this summary. One is that in most matches of *similar* targets there is not just one algorithm that fits the data but data are often fitted similarly well be several different algorithms. In some test series, even more algorithms than listed in Table 2 had produced good predictions of the data, which also reflects the high intrinsic correlation of algorithms in these tests. In matches of *dissimilar* targets, however, good fits were rare, so that the top-five listings in Tables 3 and 4 include algorithms that clearly misfit the data. Another important conclusion is that predictions from plausible combinations of presumed salience algorithms were, in general, not better than predictions from arbitrary unique algorithms. In fact, which algorithm produced the best prediction of a particular equal-salience match was not clear beforehand and could apparently not be deduced from the ranking of targets and distractors in the two patterns and their salience computation in other test series. The exact algorithms were less important for equal-saliency matches than one should expect if salience computation were based on few constant rules. Even when the data from one test series were closely fitted by a certain algorithm, this same algorithm may have failed on other test series.

Interestingly however, responses were not generally uncertain but almost all matches had revealed similar results from the different observers. This suggests an, at this stage of the project, speculative explanation. Test conditions might have been even more distinct than assumed. Instead of distinguishing DARK and BRIGHT targets in maximum and minimum configurations, it might have been important to distinguish further ranking differences like, e.g., test targets being brighter or darker than reference distractors, or, e.g., test targets being brighter than reference targets. If these further distinctions had been important but were not made, the resulting fits of various algorithms in different test series should be arbitrary and should also depend on the selection of included test conditions. There are several examples that would support such a view; the fact that equal-salience matches of homogeneous blob patterns follow different rules when backgrounds are same or different (Nothdurft, 2015); the observation that bright blob arrays are differently matched when either reference blobs or test blobs are the brightest items in the stimulus (Exp. 10; see also Nothdurft, 2015); and the performance differences seen with test series *O12* (Figs. 7 and 16) and *O12a* (Fig. 12) that was already discussed above. It might be that salience computation is based on a framework of rules, some acting more generally (like the usage of power-transforms when targets differ in lightness and are presented on largely different backgrounds), others perhaps depending on the detailed luminance distribution and item configuration in the stimulus and in a given test series.

Conclusions from discussion:

The super algorithm for luminance-defined salience does not exist. – Best fits of salience matches in simple targets are summarized:

DARK maximum targets:	algorithm 1,
BRIGHT maximum targets:	algorithm 5.

Salience variations of minimum targets are predicted from several algorithms, those of target combinations only as singular solutions of arbitrary algorithms. Cross-polarity matches are power-based.

Summary of section C

Section C has shown the following results.

1. There is no unique super algorithm that would compute target salience in all different stimulus patterns.
2. Many – even perfect – matches of experimental data were arbitrary; algorithms that closely predict the salience properties of stimulus components do not necessarily predict the matching performance when these components are compared to another.
3. Salience matches of targets that differ in contrast polarity to background (thus appearing bright or dark) are best predicted by computations on a power transform of luminance, whereas matches of targets with the same contrast polarity may also be predicted by algorithms on a linear scale.

While the latter conclusion could leave everything uncertain, the experimental observations have indicated the opposite. Almost all matches revealed similar matching results by different subjects.

GENERAL DISCUSSION

Hints for reading: This last Discussion evaluates possible methodological problems and brings a number of demonstrations that are supposed to illustrate the observed differences between DARK and BRIGHT target matches. Consequences for equal-salient target matches and the role of target location in the visual field are discussed.

When I have started the project of measuring luminance-defined salience, I expected convincing findings with clear rules (and hopefully no exceptions) that could be studied in a few months. I was wrong. Already the accompanying study on blob arrays and single targets (Nothdurft, 2015) showed that there were *several* rules (and quite a few exceptions) some even pointing into the huge field of brightness and lightness perception which I tried to avoid since it has already kept busy so many laboratories over so many years. But when moving from homogeneous arrays to popout patterns (one target, several distractors), as in the present study, the difficulties seemed to grow exponentially. Interestingly, there always seemed to be rules, as the different observers produced similar results, but these rules appeared to vary across test series and could not easily be generalized. That often suggested to expand the study and include further test series to find the general rules (with only few exceptions) behind all variations—the normal scientific progress.

The present paper is full of such expansions (some added early, some late) but has still not yet reached a final *simple and general* model, although many findings are robust and can be described by clear rules. The problem is not the eventual missing of certain rules but rather their diversity and their variability in different tests. The necessity to explain different salience matches by a large variety of complex algorithms which often can only predict the matches of one particular test combination is unsatisfactory. May be future work will provide a simpler and better synopsis of all these findings.

However, the various rules are already presented above and do not need to be repeated here. Instead, I will concentrate on two major lines of discussion. First, of course, I have to discuss potential shortcomings and pitfalls of the study. Have there been problems with generating the appropriate luminance settings in experiments? Did subjects perform the tasks as they were supposed to? Are measures real or artificial? Second, I will try to illustrate some findings of the study and discuss consequences for equal-salience matches of different targets. I will show examples of good and bad salience matches and demonstrate how salience perception may change with retinal location—an issue that is particularly important for salience-controlled gaze shifts.

Matching difficulties

Monitor luminance

Two principle problems in luminance studies are (i) eventual variations or uncertainties in measuring monitor luminance and (ii) luminance variations over the screen that occur on almost all monitors. I have tried to exclude or compensate for these problems in the present study.

Monitor luminance was repeatedly measured over the course of the study with a Spotmeter (Photo Research); in addition, the readings were two times compared (and confirmed) with readings from new or freshly calibrated Spectrometers (also Photo Research) borrowed from other labs. All readings were constant over the time of the study.

But it should be stressed that even systematic variations of photometer sensitivity or monitor luminance should not have affected the analysis of salience matches in the study. If the luminance measures *tg, dis,* and *bg* in the formulas of Table 1 had been replaced by wrong values *tg', dis',* and *bg'* which had systematically over- or underestimated the true luminance values by a factor, *f,*

$$tg' = f \cdot tg \quad , \qquad dis' = f \cdot dis \quad , \quad \text{and} \quad bg' = f \cdot bg \quad ,$$

the factor would cancel out in all predictions of section C.

Monitor inhomogeneities, and thus eventual differences between targets at different screen locations, are not easy to avoid and were also seen in the present experiments. Their effects were studied in early matching experiments. When the positions of reference and test targets on the monitor were exchanged, the resulting deviations were not larger than random variations obtained in repeated measurements with test targets on either one side alone. Since the data of the present study represent averages of matches with exchanged stimulus patterns, a possible error from monitor inhomogeneities should thus be small and cannot explain the much larger luminance variations obtained with different test conditions.

Modulation by selective attention

Notably stronger was the modulation of a target's brightness (and salience) when an observer attended to it. It is well known that attention may modulate the salience and brightness of an attended object (Carrasco, Ling, & Read, 2004; Reynolds & Desimone, 2003; Treue, 2004) in quite a similar way as luminance (and brightness) may modulate salience (this study) or attract attention (Irwin, Colcombe, Kramer, & Hahn, 2000; Spehar & Owens, 2012; Theeuwes, 1994, 1995). In a preliminary test of monitor inhomogeneities, in which an observer had to attend over many trials to one particular target location, this effect was, in fact, so dramatic that strong luminance inhomogeneities on the monitor had been assumed. However, when the monitor was turned upside-down, these differences did not rotate with the monitor but remained at the location to which the observer attended.

In the experiments of the present study, however, the effects from attentional modulation can be assumed to have been small and negligible, since observers had to compare two objects, the reference target and the test target, and were required to shift their attention frequently between these objects.

A related issue may be priming effects which can be seen when a target or stimulus pattern from a previous presentation is repeated in a subsequent presentation (Theeuwes & Van der Burg, 2013). However, since in the present study two targets had to be compared and *matched* in salience and since patterns followed each other in very slow frequency, it is unlikely that direct priming effects have played a substantial role. The fact, however, that matches in different test series could follow different algorithms (as found and discussed with Experiment 6 above) may weaken this statement. In "uncertain" matches in which targets could be matched according to different rules (as in test series *O;* cf. Figs. 7, 11 and 12), there might have been a bias to one of these rules if the matches along this rule predominated the sample.

Task difficulties

The instructions were clear; observers had to match two targets in salience. But it was not always clear, what that would mean. Although observers were instructed, to evaluate the perceived target salience and *not* any perceptual constructs such as, e.g., the apparent luminance difference of targets and distractors or the apparent sharpness (contrast gradient) of the target on background, it was not always clear what they indeed had looked at. The salience of a target is based on various aspects, on item salience and discrimination salience (as distinguished in the present study) but also on contrast and brightness (as illustrated, e.g., in Nothdurft, 2015, Fig. 10). It is likely that observers attended to all these different aspects when performing the matches but might have been biased to one or the other aspect in a certain task. An obvious example are the different matches of DARK minimum and maximum targets in Experiment 9 (Fig. 22) where subjects had sometimes ignored the low item salience of targets.

This variety of salience aspects might have been one reason why matches of some test conditions were less certain than others. But it must nevertheless be stressed that all these different aspects do contribute to the salience of a target and that there is no reasonable way to exclude one or the other aspect from the percept of salience. But it may be that some salience aspects are more important than others for a certain visual function. The quickly accessed luminance contrast of a target, for example, might be more important for the control of fast gaze shifts than the slow and careful evaluation of brightness differences. But this remains to be shown. Despite all variations, however, performance was very similar across observers and usually well reproduced in repeated runs.

Generalized findings and consequences

In the very last part of the paper I will illustrate major findings of the study and discuss some interesting and perhaps unexpected consequences. Some observations were already illustrated; the distinction of item salience and discrimination salience (Fig. 13); the different effects of background luminance on target salience (Fig. 14); and the influence of blob density on the maximum-minimum paradigm (Fig. 36).

Different algorithms for DARK and BRIGHT targets

Matches have revealed a principle difference in the salience computation of DARK and BRIGHT targets. With DARK maximum targets, equal salience was independent of background luminance and best predicted by the constant-addition rule (algorithm 1). With BRIGHT targets, however, equal salience could vary with the luminance span of background and distractors (*salmin* algorithm 5), with the constant-ratio rule (algorithm 2), or, if the target to be adjusted was the brightest item in the scene, with the constant-addition rule (algorithm 1). Thus, both target types differed in the computation of equal salience and, in addition, BRIGHT targets were less precise and less unique in the algorithm to follow.

These differences are illustrated in Figures 50 and 51[7]. In each row, the target in the reference pattern is "matched" by other targets that were predicted as equal-salient among their distractors from different rules. For *DARK targets* (Fig. 50), the overall best matches are obtained with algorithm 1. Matches from other algorithms are sometimes too little modulated (the target is less salient than in the reference pattern), e.g. algorithms 2=3 and sometimes [1], or over-modulated (the target is far more salient than in the reference pattern), e.g. algorithm 5, particularly with dimmer backgrounds (rows d and e*)*. For *BRIGHT targets* (Fig. 51), the overall best matches should be obtained with algorithms 2 and 5. With some patterns (e.g., rows b and d), better matches might lie *between* these predictions (as would predictions from the new hybrid algorithms). Some matches from algorithm 1 are too little modulated (e.g., rows b and d) and some from algorithm [1] too strongly (e.g., row c). Overall, there

7 To see all effects in this and the following figures appropriate luminance settings in the printouts are important. That might be achieved with a postscript printer that linearly reproduces the intended luminance settings; see Appendix. For inspections on a monitor you have to adjust the *gamma correction*.

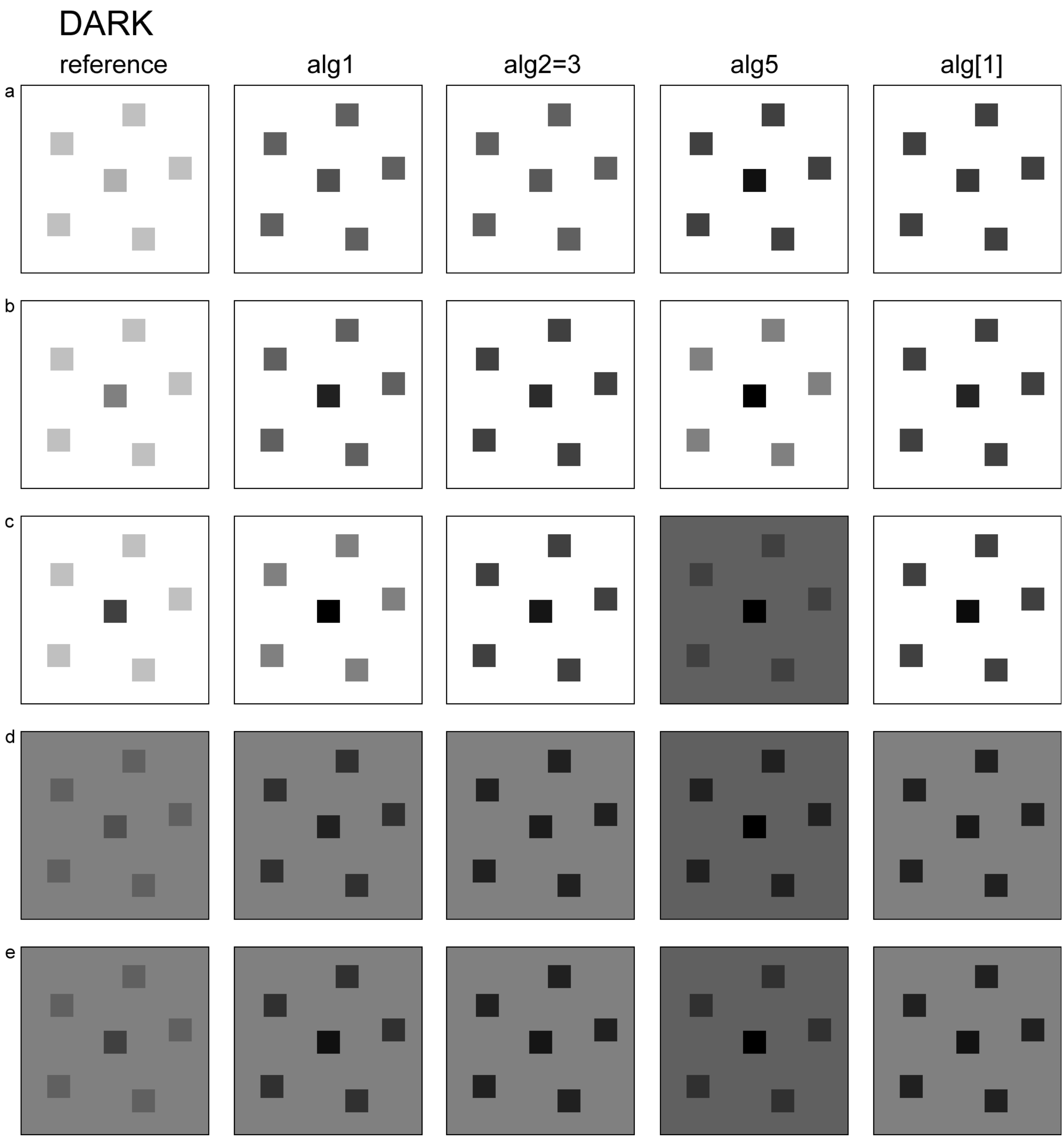


Figure 50. *Illustration of DARK maximum target matches with different algorithms.* Each row shows various matches of the reference pattern based on the algorithms listed above. The overall best matches are obtained with algorithm 1. Matches from other algorithms show often too little or too strongly salient targets. (You need a printer with linear output characteristics to see appropriate luminance variations; see Appendix.)

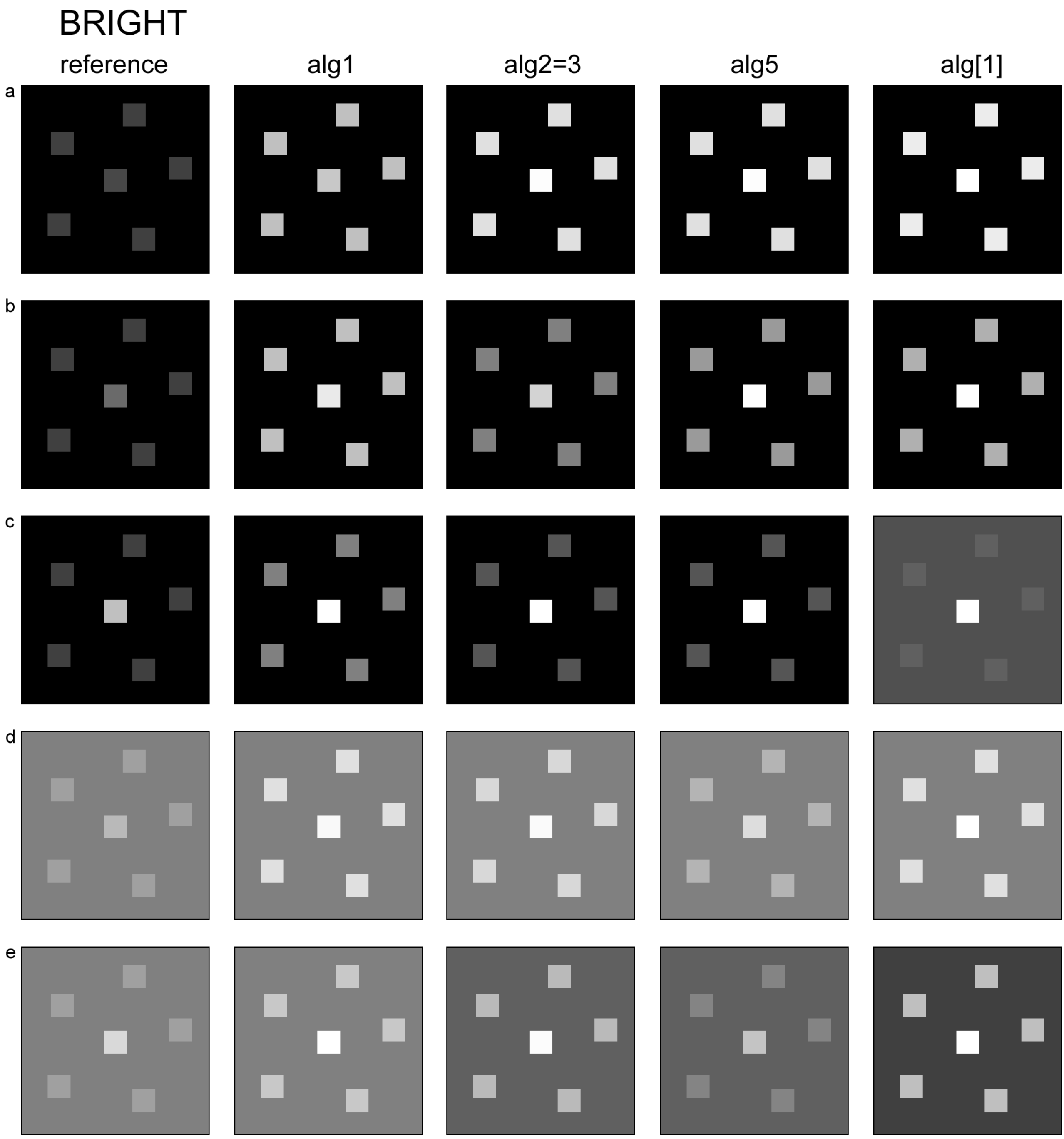


Figure 51. *Illustration of BRIGHT maximum target matches with different algorithms.* Each row shows various matches of the reference pattern based on the algorithms listed above. According to the experiments, the overall best matches should be obtained with algorithms 5 and 2 (or in between), and for particularly bright targets even with algorithm 1. In most columns some targets are slightly less or more salient than the reference target. But many targets can be accepted as about equal-salient to it. (You need a printer with linear output characteristics to see appropriate luminance variations; see Appendix.)

seem to be fewer obvious mismatches with BRIGHT targets in Figure 51 than with DARK targets in Figure 50.

Reference targets in Figure 51 were not arbitrary but computed from the DARK reference targets in Figure 50 with algorithm [1] (which gave the best results for DARK to BRIGHT cross-polarity matches). Thus it is possible to verify even these matches by comparing the salience of reference targets in Figures 50 and 51. By and large, Figures 50 and 51 thus demonstrate the main findings of the equal-salience matching experiments.

The differences between algorithms, however, are often not very strong. Some targets look equal-salient even when generated from the "wrong" algorithm. Apparently, our visual system is quite tolerant and may even accept targets as "about" equal-salient that are, in fact, notably different. (Just compare the various examples in Figs. 50 and 51.) This uncertainty in salience estimates may indicate that small variations in target luminance are less important for salience estimates than one might have thought.

Matching sequences are (theoretically) non-commutative

Given the different equal-salience algorithms for DARK and BRIGHT targets, we may go one step further. While we should be able to find two equal-salient DARK targets (algorithm 1) and to each DARK target an equal-salient BRIGHT target (algorithm [1]), the two BRIGHT targets are not necessarily equal-salient to each other (algorithm 5). That is, equal-salience matches are computationally non-commutative. Note however that convincing examples of such non-commutative matches are difficult to find in Figures 50 and 51. The large tolerance of our visual system to accept wrong predictions still as equal-salient may help to overcome the theoretically predicted misfits.

Foveal vs. peripheral inspections

An important role of salience is the guidance of eye movements to behaviorally relevant objects (Beutter, Eckstein, & Stone, 2003; Borji, Sihite, & Itti, 2013; Koehler, Guo, Zhang, & Eckstein, 2014; van Zoest & Donk, 2005). For that, salience differences of targets presented parafoveally or in the periphery of the visual field should be far more important than salience differences of targets that are already foveated, as in the experiments of the present study. Simultaneous salience matches of targets in the periphery and sequential matches of targets in the fovea may produce different results; brightness comparisons may dominate the foveal matches (Nothdurft, 2015). For many stimuli, however, the two inspection modes give similar results. Differences should mainly depend on the different spatial modulation transfer functions (MTF) at different eccentricities in the retina and thus be largely predictable (cf. Cornsweet, 1970).

The MTF has the form of an inverted U (with smaller reduction on the low spatial frequency side). When moving towards locations in retina periphery, the maximum of the MTF shifts towards lower frequencies. To illustrate and enhance the changes between foveal and peripheral inspections we may thus construct patterns that take care of these differences. An attempt is made in Figure 52. Targets and distractors in the left-hand wide blob raster have the same luminance settings as targets and distractors in the right-hand dense blob raster. Also the backgrounds are identical. When the individual targets are *foveated*, they all stand out conspicuously, except those in the bottom patterns. But if you compare their conspicuity among distractors, then some targets in the right-hand patterns stand out slightly more vividly than the corresponding targets in the left-hand patterns (since luminance differences there would be mainly represented in lower spatial frequency bands). This difference will reverse, however, when you fixate the crosses between the patterns and then look for salience differences between targets in your retina periphery. When moving your gaze from the upper to the lower crosses you may then notice that the targets in the dense blob raster disappear earlier (because the spatial frequency bands that encode the differences are strongly reduced in the MTF at this eccentricity) than the targets in the wide blob raster on the left. Variations of this sort are not specific properties of salience representation but should be similarly seen with any visual stimulus. However, these differences would be extremely important for salience-controlled eye movements, as only salience effects that are detected outside the fovea may evoke a gaze shift towards the target.

This demo illustrates the extremely sensitive competition of luminance modulation, retinal location and general sensitivity in the functional aspects of luminance-defined salience. Whether targets that differ in salience can attract attention or gaze strongly depends on their actual retinal location and hence the current gaze of the observer. Small gaze shifts can let targets appear and previously salient targets disappear from the salience map.

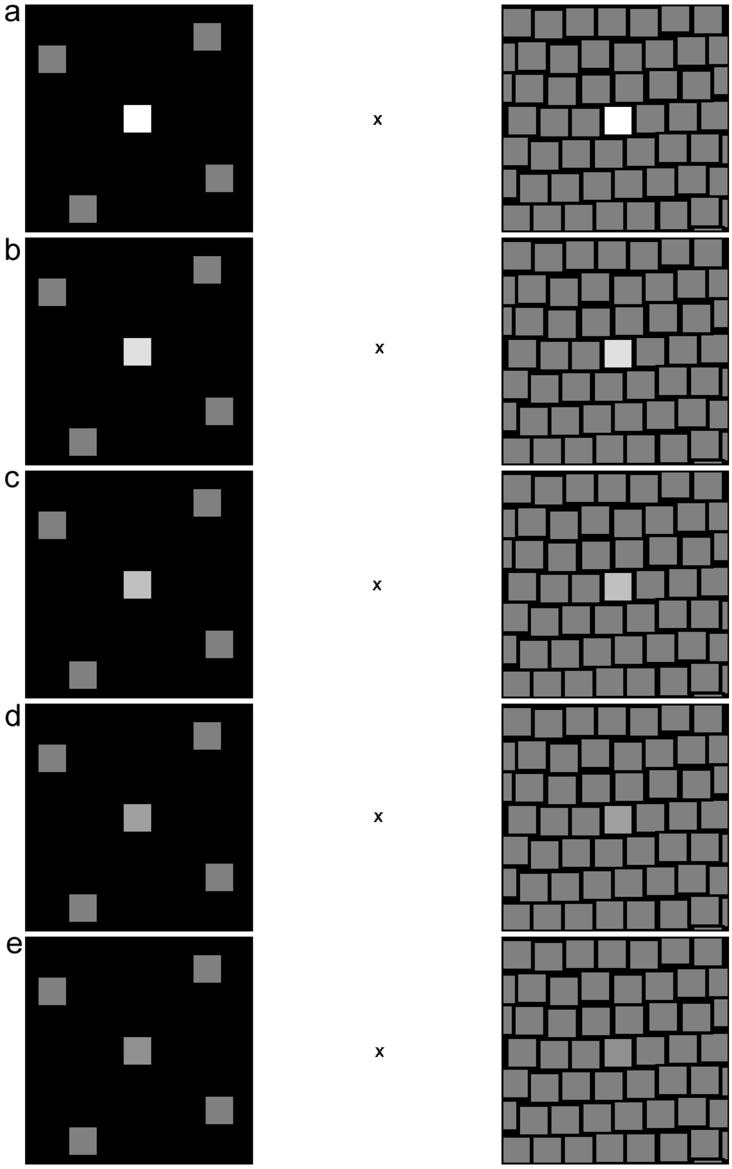


Figure 52. *Illustration of salience variations with viewing eccentricity.* The perception of luminance gradients depends on the local modulation transfer function, MTF, which varies with retinal eccentricity. This is a common phenomenon in vision and not related particularly to salience. The difference between target and distractors in the left-hand and right-hand patterns is represented in different spatial frequency bands. When foveating each target at an appropriate viewing distance, some targets in the left-hand patterns may appear slightly less conspicuous than the according targets in the right-hand patterns, e.g., rows *b* and *c*. When looking at the targets parafoveally or in the retina periphery, e.g. by fixating the crosses in the middle, the ranking may change and some targets in the left-hand patterns may look more conspicuous than the according targets in the right-hand patterns, e.g., row *d*. Luminance settings in left- and right-hand patterns of a row are identical. (You need a printer without local contrast enhancement; see Appendix.)

This effect is particularly dramatic if targets are embedded in (many) distractors, as illustrated in Figure 53. The two targets in the upper picture are made equal-salient (algorithm 1) in their surroundings (which you may verify by foveating the targets in alternation). But which target would attract your gaze depends on where you actually are looking (beside many other effects). If you happen to look at blobs on the left- or right-hand side, the nearer target appears more salient and will likely attract your gaze and attention. You may verify this by fixating various locations on the upper or lower frame. Only when you happen to look at blobs in the middle near the central cross, should both targets be equal-salient and equally attract your gaze.

There are various ways to make such targets more attractive. One is to reduce the blob density and minimize background structure. As can be experienced in the lower picture of Figure 53, removing all distractors strongly enhances the conspicuity of *both* targets for fixations everywhere in the pattern. Both targets may now attract your gaze and attention, although they clearly differ in salience. Their different saliency, however, is only noticed when you fixate the central cross. Another way to increase a target's salience would be to increase its contrast and thus make it more salient than before. But when you compare the two examples in Figure 53 you will notice that these targets are already quite salient. In this example, the removal of distractors to overcome restrictions from the limited spatial resolution in retina periphery has a much stronger effect on the target's conspicuity (outside the fovea) than gradual increments of the target contrast.

Our natural visual environment is highly structured and full of gradual luminance variations; luminance-defined salience effects in such a surrounding should strongly suffer from the modulation transfer function at different retinal locations. It may thus be that luminance-defined salience differences play only a minor functional role in such surroundings except when targets are large or spatially isolated, or particularly bright or dark. The interaction of luminance modulation and positional effects on the retina may also explain why some studies did, and others failed to find a strong

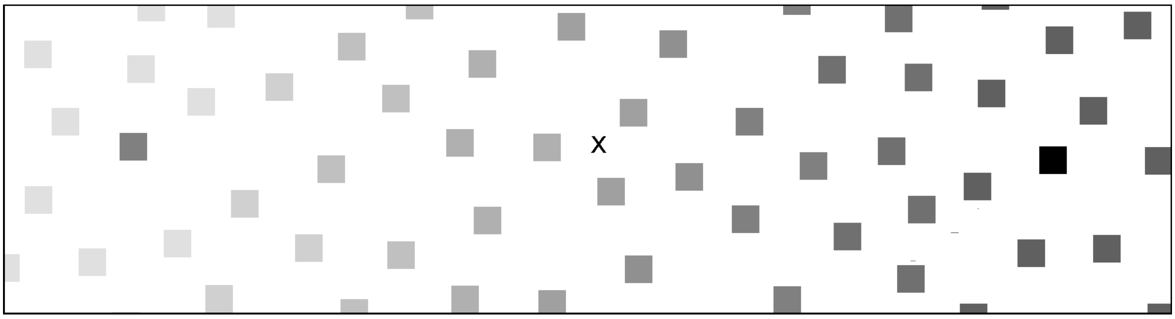
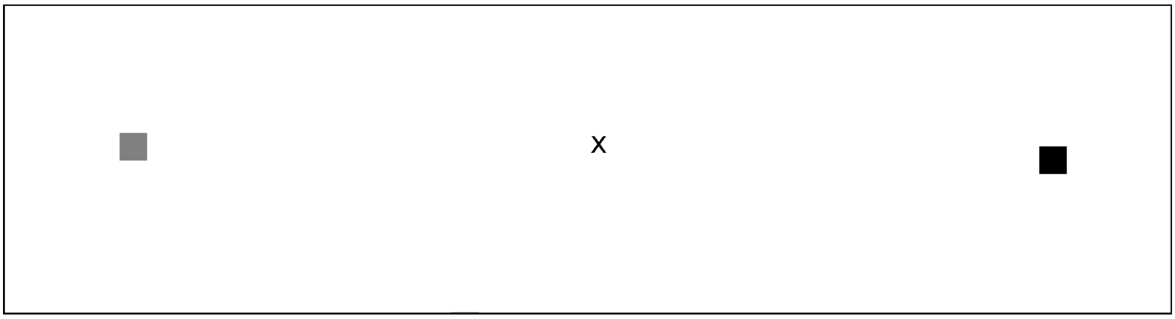


Figure 53. *Luminance-defined salience depends on contrast, structure, and eccentricity.* In the pattern above, the two targets are equal-salient in their surroundings. But which of them attracts your gaze depends on where you are looking. Try it out by fixating various locations on the upper or lower frame. Only when you look at the central cross, both targets are equal-salient and may equally attract your gaze. When distractors are removed, as in the pattern below, both targets may attract your gaze from all over the frame, despite their different item salience. (You may need a postscript printer to see the effect; see Appendix.)



correlation of luminance-defined salience and evoked gaze shifts (Borji, Sihite, & Itti, 2013; Einhäuser and König, 2003).

ACKNOWLEGDEMENT

I thank all subjects for patient collaboration and Barry Lee for instrumental support and helpful discussions about luminance photometry.

REFERENCES

Anderson, B.L., Singh, M., & Meng, J. (2006). The perceived transmittance of inhomogeneous surfaces and media. *Vision Research, 46,* 1982-1995.

Arend, L.E., & Goldstein, R. (1987). Simultaneous constancy, lightness and brightness. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 4(12),* 2281-2285.

Beutter, B.R., Eckstein, M.P., & Stone, L.S. (2003). Saccadic and perceptual performance in visual search tasks. I. Contrast detection and discrimination. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 20(7),* 1341-1351.

Blakeslee, B., & McCourt, M.E. (2013). Brightness induction magnitude declines with increasing distance from the inducing field edge. *Vision Research, 78,* 39-45.

Borji, A., Sihite, D.N., & Itti, L. (2013). What stands out in a scene? A study of human explicit saliency judgment. *Vision Research 91,* 62-77. http://dx.doi.org/10.1016/j.visres.2013.07.016

Braun, J. (1994). Visual search among items of different salience: removal of visual attention mimics a lesion in extrastriate area V4. *Journal of Neuroscience, 14,* 554-567.

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience, 7(3),* 308-313.

Cornsweet, T.N. (1970). Visual perception. Academic Press Inc., New York & London.

Dannemiller, J.L., & Stephens, B.R. (2001). Asymmetries in contrast polarity processing in young human infants. *Journal of Vision 1(2):5,* 112-125. http://journalofvision.org/1/2/5/

Einhäuser, W., & König, P. (2003). Does luminance-contrast contribute to a saliency map for overt visual attention? *European Journal of Neuroscience, 18,* 1089-1097.

Gilchrist, A. (1988). Lightness contrast and failures of constancy: a common explanation. *Perception & Psychophysics, 43(5),* 415-424.

Irwin, D.E., Colcombe, A.M., Kramer, A.F., Hahn, S. (2000). Attentional and oculomotor capture by onset, luminance and color singletons. *Vision Research, 40,* 1443–1458.

Kingdom, F.A.A. (2011). Lightness, brightness and transparency: a quarter century of new ideas, captivating demonstrations and unrelenting controversy. *Vision Research, 51,* 652-673.

Koehler, K., Guo, F., Zhang, S., & Eckstein, M.P. (2014). What do saliency models predict? *Journal of Vision, 14(3):14,* 1–27. http://www.journalofvision.org/content/14/3/14, doi:10.1167/14.3.14

Maertens, M., & Wichmann, F. A. (2013). When luminance increment thresholds depend on apparent lightness. Journal of Vision, 13(6):21, 1–11, http://www.journalofvision.org/content/13/6/21, doi:10.1167/13.6.21.

Moore, C.M., & Brown, L.E. (2001). Preconstancy information can influence visual search: the case of lightness constancy. *Journal of Experimental Psychology: Human Perception and Performance 27(1),* 178-194.

Moulden, B., Kingdom, F., & Gatley, L.F. (1990). The standard deviation of luminance as a metric for contrast in random-dot images. *Perception 19 (1),* 79-101.

Nothdurft, H.C. (2000). Salience from feature contrast: variations with texture density. *Vision Research, 40,* 3181-3200.

Nothdurft, H.C. (2002). Attention shifts to salient targets. *Vision Research, 42,* 1287-1306.

Nothdurft, H.C. (2006a). Salience-controlled visual search: are the brightest and the least bright targets found by different processes? *Visual Cognition, 13(6),* 700-732.

Nothdurft, H.C. (2006b). Salience and target selection in visual search. *Visual Cognition, 14(4-8),* 514-542.

Nothdurft, H.C. (2015). Luminance-defined salience of homogeneous blob arrays. *VPL-reports, 1*, 1-38. www.vpl-reports.de/1/

Peli, E., (1990). Contrast in complex images. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 7(10),* 2032-2040.

Reynolds, J. H., & Desimone, R. (2003). Interacting roles of attention and visual salience in V4. *Neuron, 37,* 853-863.

Robilotto, R., & Zaidi, Q., (2006). Lightness identification of patterned three-dimensional, real objects. *Journal of Vision, 6(1):3,* 18-36. http://journalofvision.org/6/1/3/

Rudd, M.E., & Popa, D. (2007). Stevens's brightness law, contrast gain control, and edge integration in achromatic color perception: a unified model. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 24(9),* 2766-2782. Erratum in 24(10), 3335.

Schiller, P.H. (1993). The effects of V4 and middle temporal (MT) area lesions on visual performance in the rhesus monkey. *Visual Neuroscience, 10,* 717-746.

Schiller, P.H., & Lee, K. (1991). The role of the primate extrastriate area V4 in vision. *Science, 251,* 1251-1253.

Schirillo, J.A. (1999a). Surround articulation. I. Brightness judgments. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 16,* 793-803.

Schirillo, J.A. (1999b). Surround articulation. II. Lightness judgments. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 16,* 804-811.

Singh, M., & Anderson, B.L. (2006). Photometric determinants of perceived transparency. *Vision Research 46,* 879-894.

Spehar, B., and Owens, C. (2012). When do luminance changes capture attention? *Attention Perception Psychophysics, 74,* 674-690. doi:10.3758/s13414-011-0266-8

Stevens, S.S. (1961). To honor Fechner and repeal his law. *Science 133,* 80-86.

Theeuwes, J. (1994). Stimulus-driven capture and attentional set: Selective search for color and visual abrupt onsets. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 799-806.

Theeuwes, J. (1995). Abrupt luminance change pops out; abrupt color change does not. *Perception & Psychophysics, 57(5),* 637-44.

Theeuwes, J., & Van der Burg, E. (2013). Priming makes a stimulus more salient. *Journal of Vision, 13(3):21,* 1–11. http://www.journalofvision.org/content/13/3/21, doi:10.1167/13.3.21

Töllner, T., Zehetleitner, M., Gramann, K., Müller, H.J. (2011). Stimulus saliency modulates pre-attentive processing speed in human visual cortex. PloS ONE 6(1): e16276. .doi:10.1371/journal.pone.0016276

Treue, S. (2003). Visual attention: the where, what, how and why of saliency. *Current Opinion in Neurobiology, 13,* 428-432.

Treue, S. (2004). Perceptual enhancement of contrast by attention. *Trends in Cognitive Sciences, 8(10),* 435-437.

Turatto, M., & Galfano, G. (2000). Color, form and luminance capture attention in visual search. *Vision Research, 40,* 1639-1643.

van Zoest, W., & Donk, M. (2005). The effects of salience on saccadic target selection. *Visual Cognition, 12(2),* 353-375.

Zehetleitner, M., Koch, A.I., Goschy, H., & Müller, H.J. (2013). Salience-based selection: attentional capture by distractors less salient than the target. *PloS ONE 8(1): e52595.* doi:10.1371/journal.pone.0052595

Zenger-Landolt, B., & Fahle, M. (2001). Discriminating contrast discontinuities: asymmetries, dipper functions, and perceptual learning. *Vision Research, 41,* 3009-3021.

Table 1. *Salience algorithms used in predictions.*

For each algorithm, the proposed salience of the reference target was calculated from its luminance settings in a given test condition. It was equated with the matched salience of the test target, the luminance settings of which could then be predicted from the known background and distractor settings. Quadratic deviations of predicted from experimental results were calculated and averaged over all test conditions of a particular test series to obtain the *mean squared deviation, MSD,* for a given algorithm and test series block. See section C for more details.

Reference in text		*proposed salience algorithm*	*Description, Comments*
#	*name*		
Target-Distractor Comparisons			
1	Luminance Difference	$sal \sim \lvert tg - dis \rvert$	luminance difference of target and distractors
	Luminance Ratio	$sal \sim tg/dis$	ratio of target and distractor luminance; predictions identical to those from (2)
2	Weber Contrast	$sal \sim \frac{\lvert tg - dis \rvert}{dis}$	Weber Contrast of target and distractors
2a	Whittle Contrast	$sal \sim \frac{\lvert tg - dis \rvert}{\min(dis, tg)}$	Whittle's form of Weber Contrast; ≡(2) for tg>dis
2b	max	$sal \sim \frac{\lvert tg - dis \rvert}{\max(dis, tg)}$	luminance difference normalized to maximum; ≡(2) for dis>tg
3	Michelson Contrast	$sal \sim \frac{\lvert tg - dis \rvert}{tg + dis}$	Michelson Contrast of targets and distractors
Target-Distractor Luminance Differences in Various Normalizations Including Background Luminance Settings			
4	"sal-bg"	$sal \sim \frac{\lvert tg - dis \rvert}{bg}$	normalization to background luminance
5	"salmin"	$sal \sim \frac{\lvert tg - dis \rvert}{\lvert dis - bg \rvert}$	normalization to the background-distractor luminance span
6	Diff | lum-range	$sal \sim \frac{\lvert tg - dis \rvert}{\max - \min}$	Normalization to full luminance range; for minimum target configurations ≡(5)
7	Diff | max-lum	$sal \sim \frac{\lvert tg - dis \rvert}{\max}$	normalization to luminance maximum
8	"Singh"	$sal \sim \frac{\frac{\lvert tg - dis \rvert}{tg + dis}}{\max - \min}$	Michelson Contrast of target and distractors, normalized to the full luminance-range in the pattern
8a	"Singh-spec"	$sal \sim \frac{\frac{\lvert tg - dis \rvert}{tg + dis}}{\lvert bg - dis \rvert}$	Michelson Contrast of target and distractors, normalized to the background-distractor luminance span
8b	"Weber-Singh"	$sal \sim \frac{\frac{\lvert tg - dis \rvert}{dis}}{\max - \min}$	Weber Contrast of target and distractors, normalized to the full luminance-range in the pattern

ctd.

Table 1 (ctd.)

Intermediate Ratio Computations			
	Ratio of Relative Differences	$sal \sim \mid tg - dis \mid \; : \; \mid dis - bg \mid$	ratio of the target-distractor and distractor-background luminance differences; ≡(5)
9	Ratio of Relative WC WCtg:dis:WCdis:bg	$sal \sim \frac{\mid tg - dis \mid}{dis} : \frac{\mid dis - bg \mid}{bg}$	ratio of the target-distractor and distractor-background Weber Contrasts
10	Ratio of Relative MC MCtg:dis:MCdis:bg	$sal \sim \frac{\mid tg - dis \mid}{tg + dis} : \frac{\mid dis - bg \mid}{dis + bg}$	ratio of the target-distractor and distractor-background Michelson Contrasts
Difference and Ratio Computations of Item Contrasts (to Background)			
	Ratio of Absolute Differences	$sal \sim \mid tg - bg \mid \; : \; \mid dis - bg \mid$	ratio of the target-background and distractor-background luminance differences; ≡(11)
11	Ratio of Absolute WC	$sal \sim \frac{\mid tg - bg \mid}{bg} : \frac{\mid dis - bg \mid}{bg}$	ratio of the target-background and distractor-background Weber Contrasts
12	Ratio of Absolute MC	$sal \sim \frac{\mid tg - bg \mid}{tg + bg} : \frac{\mid dis - bg \mid}{dis + bg}$	ratio of the target-background and distractor-background Michelson Contrasts
Differences of Item Contrasts (to Background)			
	WCtg-WCdis	$sal \sim \frac{\mid tg - bg \mid}{bg} - \frac{\mid dis - bg \mid}{bg}$	difference of the target-background and distractor-background Weber Contrasts; ≡(4)
13	MCtg-MCdis	$sal \sim \frac{\mid tg - bg \mid}{tg + bg} - \frac{\mid dis - bg \mid}{dis + bg}$	difference of the target-background and distractor-background Michelson Contrasts
14	WC (bg)	$sal \sim \frac{\mid tg - bg \mid}{bg}$	Weber Contrast of target to background (no influence from distractor luminance)
15	MC (bg)	$sal \sim \frac{\mid tg - bg \mid}{tg + bg}$	Michelson Contrast of target to background (no influence from distractor luminance)
Further Intermediate Computations			
A	Ratio of Relative to Absolute WC WCtg:dis:WCtg:bg	$sal \sim \frac{\mid tg - dis \mid}{dis} : \frac{\mid tg - bg \mid}{bg}$	ratio of the target-distractor and target-background Weber Contrasts
B	Difference of Absolute and Relative WC WCtg:dis-WCtg:bg	$sal \sim \left\lvert \frac{\mid tg - bg \mid}{bg} - \frac{\mid tg - dis \mid}{dis} \right\rvert$	difference of the target-background and the target-distractor Weber Contrasts
C	Ratio of Relative to Absolute MC MCtg:dis:MCtg:bg	$sal \sim \frac{\mid tg - dis \mid}{tg + dis} : \frac{\mid tg - bg \mid}{tg + bg}$	ratio of the target-distractor and target-background Michelson Contrasts
D	Difference of Absolute and Relative MC MCtg:dis:MCtg:bg	$sal \sim \left\lvert \frac{\mid tg - bg \mid}{tg + bg} - \frac{\mid tg - dis \mid}{tg + dis} \right\rvert$	difference of the target-background and the target-distractor Michelson Contrasts
E	Special Ratio of Differences	$sal \sim \frac{\mid tg - dis \mid}{\mid tg - bg \mid}$	ratio of target-distractor and target-background luminance differences; for maximum targets ≡(6); predictions identical to those from (11)

Table 2. *Best MSD values of similar target matches (Exp. 1-4, 7-8).*

For every target type (rows) and test series block (columns), the table lists the five smallest MSD values of predicted and measured equal salience matches, together with the according algorithm (gray cells) from Table 1. Best values in each cell are printed red. MSD values were computed for all test conditions of a given test, including wide and dense blob configurations. Rows labeled ”power|0.33” show the MSD values for predictions based on power-transforms of luminance (exponent $x=0.33$); for details see text in section C. Note that algorithms can be identical (symbol ≡) for a given target type or may lead to identical predictions (symbol =) even when the according salience measures differ. General identities are listed in the first column; they are valid for the entire row and the accompanying power|0.33 row underneath. Algorithms that produced the same result in a given test condition are listed together, with that value. Note that MSD values in a cell may differ considerably, indicating a clear distinction of better and worse predictions from different algorithms, or may be generally small (or large), indicating a generally good (or bad) performance of several algorithms in predicting the experimental data of this particular test condition.

		(1) Series K		(2) Series L/O		(3) Series F		(4) Series J		(5) **total**	
		alg		alg		alg		alg		alg	
	DARK Targets										
	Maximum Conditions	13	2,3	8	3,3	2	1,2	1=4	8,5	1	5,4
	(2≡2a=2b=3)	2	2,8	13	13,0	1	1,4	8	9,5	13	9,6
(A)	*(4≡7)*	1	3,0	8b=A	13,4	13	1,7	13	16,1	4	11,3
	(5≡11=6=E)	A	13,3	2	15,8	A	7,9	8b=A	30,9	A	19,6
	(14=15)	4	16,0	14	460,7	C	13,1	2	52,3	8	20,1
	power|0.33	[4]	1,7	[1,4]	12,2	[4]	0,6	[1,4]	19,6	[4]	10,6
		[13]	2,5	[13]	15,4	[13]	0,8	[8]	33,8	[1]	10,7
(B)		[2]	2,8	[8]	15,6	[2]	1,2	[13]	46,7	[13]	22,7
		[1]	2,9	[2]	15,8	[1]	1,3	[2]	52,3	[2]	25,3
		[A]	27,0	[14]	460,7	[A]	17,4	[8b,A]	63,5	[8]	28,7
	Minimum Conditions	A	0,8	10	2,3	A	0,6			10	1,4
	(2≡2a=2b=3)	10	0,9	A	3,9	13	0,9			A	1,8
(C)	*(4≡7) (8≡8a)*	C	1,3	C	7,1	10	1,1			C	3,8
	(5≡11=6=E)	12	2,0	12	12,7	B	2,6			12	6,7
	(14=15)	5	2,9	5	27,9	9	2,8			5	12,6
	power|0.33	[A]	1,0	[A]	2,7	[9]	0,6			[A]	1,8
		[10]	1,2	[10]	4,8	[4]	0,8			[10]	2,8
(D)		[12,C]	1,3	[C]	5,7	[B]	1,6			[C]	3,2
		[5]	1,7	[12]	6,2	[A]	2,1			[12]	3,5
		[8b]	1,8	[5]	10,8	[10]	2,9			[8]	5,0
	BRIGHT Targets										
	Maximum Conditions	8b	6,0	2	15,5	8b	13,9	5	4,9	5	21,4
	(2≡2a=3=2b≡7)	5	6,1	1,4	83,1	5	14,2	2	5,7	2	28,0
	(5≡11=6=E)	8a	8,9	14	401,8	8a	17,7	C	6,1	C	32,3
(E)	*(14=15)*	8	11,6			8	20,2	10	9,8	13	35,4
		13	13,9	Series LX:		2	24,0	13	14,5	10	44,7
				2	16,2						
	power|0.33	[8b]	18,1	[13]	12,0	[13]	14,1	[13]	4,8	[13]	18,3
		[5]	18,2	[2]	15,5	[1]	23,8	[2]	5,7	[2]	28,0
		[8a]	19,9	[1,4]	31,1	[2]	24,0	[5]	8,6	[1]	34,9
(F)		[8]	20,5	[14]	401,8	[8b]	29,4	[C]	11,5	[5]	43,0
		[C]	25,4	Series LX:		[5]	29,5	[10]	12,4	[8b]	47,2
				[2]	16,2						
	Minimum Conditions	5,8b	1,7	5,10	1,6	8	3,0			5	2,2
	(2≡2b≡7=2a=3)	10	2,7	C	2,0	8b	4,9			10	3,6
(G)	*(5≡11=6=E)*	C	3,9	12	3,3	5	5,0			C	5,0
	(8≡8a)	8	5,6	13	3,9	10	12,4			8	10,7
	(14=15)	12	8,3	A	4,4	1	16,0			8b	12,0
	power|0.33	[5,8b]	3,1	[5]	1,5	[13]	8,6			[5]	4,4
		[8]	3,2	[10]	1,7	[8]	15,1			[8]	5,5
(H)		[10]	4,3	[C]	1,8	[1]	16,4			[10]	6,1
		[C]	4,5	[12]	1,9	[2]	16,5			[C]	6,3
		[12]	5,0	[A]	2,4	[8b]	16,8			[8b]	6,9

For entries in cell *E4* (BRIGHT maximum targets, Series J) the first test conditions with test targets brighter than reference targets are not included in analysis

Table 2H. *Improvements by the use of hybrid algorithms.*

In addition to Table 2 rows E and F, the table lists the MSD values for hybrid algorithms. Best values are printed in red and are eventually copied from Table 2 if there was a better MSD value than obtained here.

		(1)		(2)		(3)		(4)		(5)	
		Series K		Series L/O		Series F		Series J		**total**	
		alg		alg		alg		alg		alg	
	BRIGHT Targets *Maximum Conditions*										
	Hybrid Algorithms	8b	6,0	2	15,5	<1,5>	4,5	<2,5>	4,6	<2,5>	13,1
	vs. Best Predictions	<1,5>	10,7	<1,5>	16,4	<2,5>	4,6	<1,5>	20,4	<1,5>	16,5
	Otherwise	<2,5>	11,1	<1,2>	34,4	<1,2>	23,7	<1,2>	27,6	<1,2>	39,5
(E)		<1,2>	66,4	<2,5>	-						
				Series LX:							
				2	16,2						
				<2,5>	21,3						
				<1,5>	39,4						
				<1,2>	60,9						
	power|0.33	<[1],[5]>	6,8	[13]	12,0	<[2],[5]>	6,0	[13]	4,8	<[1],[5]>	12,3
		<[2],[5]>	6,9	<[1],[2]>	25,8	<[1],[5]>	6,0	<[2],[5]>	4,8	<[2],[5]>	14,2
		<[1],[2]>	67,0	<[1],[5]>	-	<[1],[2]>	23,9	<[1],[5]>	5,2	<[1],[2]>	31,2
(F)				<[2],[5]>	-			<[1],[2]>	9,9		
				Series LX:							
				[2]	16,2						
				<[1],[5]>	17,0						
				<[1],[2]>	21,1						
				<[2],[5]>	28,2						

Table 3. *Best MSD values for minimum to maximum target matches in Experiments 9-10, 12 and 13.*

Presentation as in Table 2, except that predictions were obtained from "mixed combinations" and "unique algorithms" (see text). Computations were made on all test series in a given block, including matches from dense and wide configurations. However, since these two configurations had produced quite different matches for DARK targets in test series block *J*, these data were also split for an additional analysis (cf. Table 3a).

		(1)		(2)		(3)		(4)	
		Series J				Series WM1, WM4			
		mixed		unique		mixed		unique	
		alg		alg		alg		alg	
	Minimum vs. Maximum Conditions								
	DARK Targets	5:A	78.8 *	9	59.2 *	A:1	34,6	3=8a=10	8,3
	identities (unique algorithms only):	A:1	88.2 *	8a	92.9 *	5:A	35,7	8	37,6
(A)	*(1=4≡7) (2a=2b=3)*	10:A	113.1 *	10	128.1 *	C:1	43,6	11=12=14	905,4
	(14=15)	C:1	137.2 *	8	128.8 *	10:1	76,6		
		8:13	213.3 *	5	134.9 *	5:1	78,0		
	power|0.33	[8:2]	49.6 *	[9]	93.9 *	[8:4]	6,1	[3=8a=10]	8,3
		[8:4]	52.3 *	[8]	114.3 *	[8:2]	10,2	[13]	9,3
(B)		[5:A]	115.2 *	[8a]	115.9 *	[8:13]	17,3	[1=2=5=9]	19,0
		[8:1]	133.8 *	[10]	131.3 *	[5:1]	31,2	[11=12=14]	905,4
		[10:A]	144.4 *	[5]	132.5 *	[10:1]	34,3		
	BRIGHT Targets	5:10	5,0	10	9,0	10:2	9,4	1=2=5=9	9,6
	identities (unique algorithms only):	C:5	16,0	2	52,5	5:2	11,3	11=12=14	630,8
(C)	*(1=4) (2a=3=2b≡7)*	C:2	34,7	5	77,4	10:5	19,3		
	(14=15)	5:2	39,1	1	84,6	C:5	56,3		
		10:2	49,8	8a	105,3	8b:13	126,2		
	power|0.33	[5:10]	9,6	[2]	2,1	[8b:2]	12,3	[11=12=14]	630,8
		[8:13]	13,5	[10]	18,3	[8b:C]	27,2		
(D)		[8b:2]	47,9	[5]	40,9	[8b:10]	41,7		
		[10:5]	51,2	[8a]	44,3	[8:2]	44,4		
		[8:2]	101,1	[1]	54,2	[8b:5]	67,3		

* Large differences between MSD values for wide and dense arrangements due to item salience effects. Individual values are shown in Table3a.

Table 3 (ctd)

		(5)		(6)		(7)		(8)	
		Series WZ				**total**			
		mixed		unique		mixed		unique	
		alg		alg		alg		alg	
	Minimum vs. Maximum Conditions								
	DARK Targets	5:A	49,6	8a	29,8	5:A	70,6	9	64,2
	identities (unique algorithms only):	C:1	99,5	5	44,6	A:1	86,9	8a	75,6
(A)	*(1=4≡7) (2a=2b=3)*	A:1	147,5	10	46,3	10:A	112,7	10	104,2
	(14=15)	8:13	152,6	9	132,1	C:1	120,9	5	117,0
		10:1	156,1	14	172,6	8:13	189,4	8	118,2
	power|0.33	[8:4]	41,1	[9]	29,7	[8:2]	43,5	[9]	77,8
		[8:2]	41,2	[8a]	39,4	[8:4]	44,9	[8a]	94,1
(B)		[8:1]	89,5	[5]	46,7	[5:A]	111,4	[10]	106,8
		[12:1]	98,5	[10]	48,0	[8:1]	118,1	[5]	109,1
		[5:1]	134,9	[14]	172,6	[5:1]	143,0	[8]	118,9
	BRIGHT Targets	C:5	22,4	2	31,4	C:5	27,2	2	37,5
	identities (unique algorithms only):	5:2	25,4	10	53,6	5:2	29,4	10	40,8
(C)	*(1=4) (2a=3=2b≡7)*	10:2	28,7	5	73,0	10:2	35,4	5	60,0
	(14=15)	10:5	86,8	8	101,8	C:2	53,3	1	99,4
		5:10	115,9	8a	107,9	5:10	53,5	3	101,4
	power|0.33	[8:13]	11,7	[2]	25,2	[8:13]	30,0	[2]	12,1
		[C:5]	31,2	[8b]	50,5	[8b:2]	37,1	[5]	41,7
(D)		[8b:2]	38,0	[10]	56,5	[5:10]	38,6	[10]	46,6
		[5:10]	65,7	[5]	61,3	[10:5]	46,4	[8b]	57,6
		[10:5]	67,0	[8a]	71,8	[8:2]	86,7	[8a]	64,5

Table 3a. *Split analysis for DARK targets in series J of Table 3.*

Analysis was performed on different sub-samples of the data, *dense, wide sample1,* and *wide sample2* (see Fig.23). Only the three (four) best values are given in each cell; most other values were above 100.

		(1a)		(1b)		(1c)		(2a)		(2b)		(2c)	
		Series J											
		mixed						unique					
		dense		wide sample1		wide sample2		dense		wide sample1		wide sample2	
		alg		alg		alg		alg		alg		alg	
	Minimum vs. Maximum Conditions												
	DARK Targets	10:A	38,1	10:A	36,2	C:1	32,8	9	15,5	9	22,0	8a	5,4
	(identities as in Table 3)	5:A	48,4	5:A	89,2	8:13	37,2	1	86,7	1	61,0	10	8,5
(A)		A:1	69,4	A:1	113,3	10:1	39,9	6	96,9	8	61,3	5	12,4
	power|0.33	[8:4]	18,3	[8:4]	36,4	[8:1]	10,2	[8]	79,2	[8]	52,6	[8a,9]	7,0
		[8:2]	18,6	[8:2]	39,2	[8:A]	38,5	[6]	81,9	[6]	82,0	[10]	9,5
(B)		[5:A]	47,9	[10:A]	62,1	[8:2]	91,1	[9]	108,3	[1]	99,9	[5]	10,4
		[10:A]	62,2	[5:A]	62,6								

Table 4. *Best MSD values for the cross-polarity matches in Experiment 11.*

		(1)		(2)		(3)		(4)	
		Series <W,WB1>				Series WB2			
		different		unique		different		unique	
		alg	dense	alg	dense	alg	dense	alg	dense
	DARK vs. BRIGHT Targets					**BRIGHT vs. DARK Targets**			
	Maximum Conditions	A:2	23,9	3	39,5	13:4	13,2	3	5,7
	identities (unique):	2:10	26,3	1	72,8	C:2	14,2	13	9,2
(A)	(2a=2b=3)	A:5	60,4	2	75,5	10:A	23,3	8	11,8
	(5=6=11=E)	4:C	78,8	5	76,1	C:A	25,1	C	13,5
		13:C	84,2	8b	79,8	13:2	42,4	10	19,0
	power|0.33	[A:1]	15,3	[1]	29,8	[8b:4]	7,1	[13]	4,0
		[A:5]	22,8	[2]	30,1	[13:4]	9,6	[3]	5,7
(B)		[1:5]	24,1	[7]	35,4	[8b:2]	10,1	[7]	11,4
		[4:2]	46,4	[13]	39,3	[13:8]	10,5	[5,8b]	11,6
		[13:8b]	58,1	[3]	39,5	[C:1]	15,3	[1]	12,3

APPENDIX

To see the intended effects in Figures 13, 14, and 50-53 your printer should produce linear luminance variations, as for example obtained with a postscript printer. In particular, there should be no contrast enhancement in the print. If you have a possibility to measure the luminance of illuminated paper, print out Figure A1 and measure the reflected light in the individual patches. It is important that the illumination does not vary during these measurements,

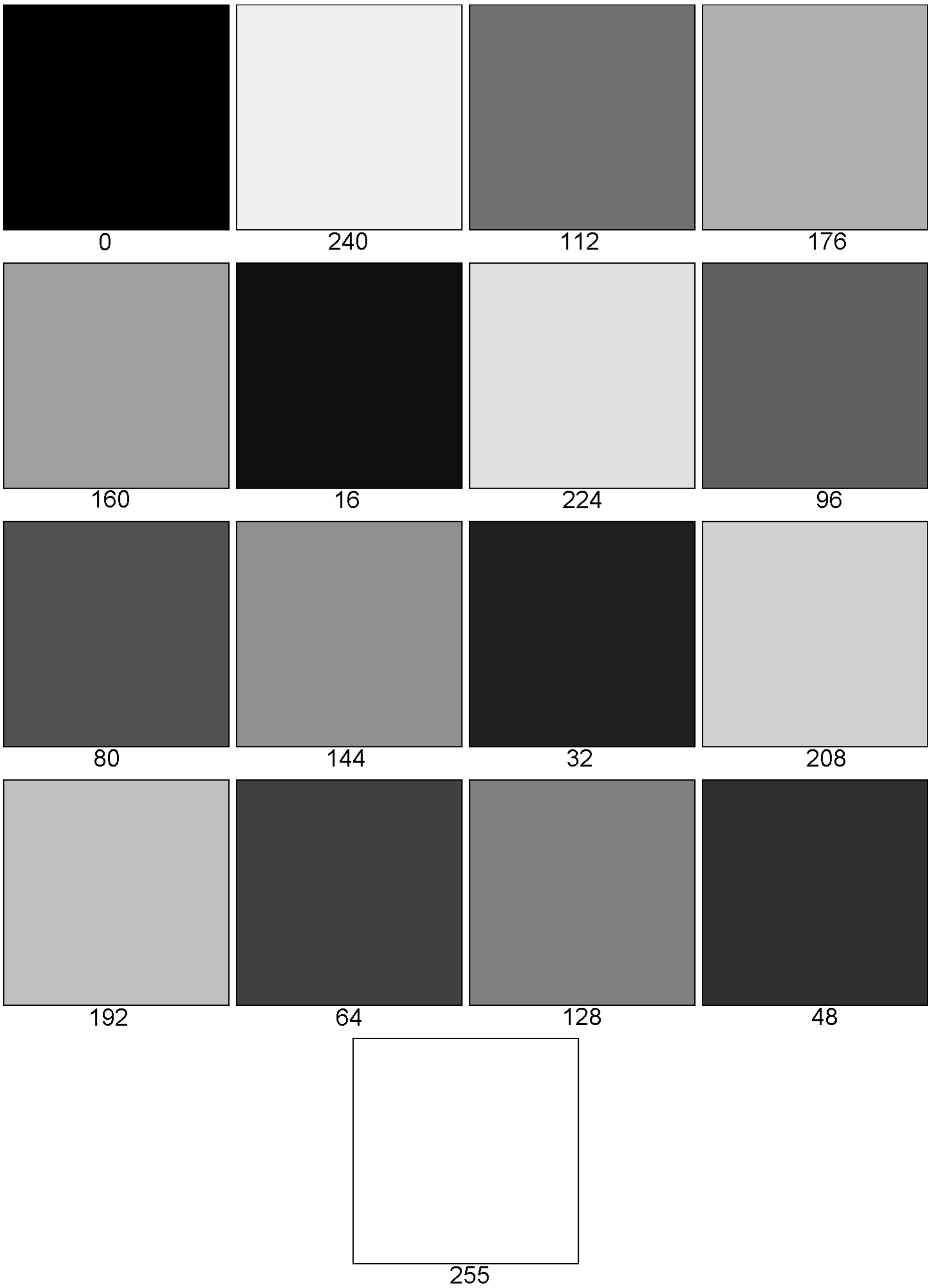


Figure A1. *Linear printout scheme.* If you have the possibility to measure the luminance of illuminated paper, print this figure and measure the luminance of the various squares under constant illumination (hold the setup stable and only move the printout). Ideally, you should obtain a linear variation of the measures with the numbers given under each square.

so hold the setup constant and move the paper, not the lights. If you get a nearly linear relationship between the measured luminance and the values printed underneath the various patches, you should have a good chance to see the above figures as intended. Using printouts of this kind (with slightly different luminance settings) the author has obtained a fairly linear relation of luminance settings and measured luminance for printouts from a postscript printer. Note that linear luminance variations in Figure A1 do *not* represent linear sensation differences in perception.

To see all intended effects on a *monitor*, however, you likely have to modify the gamma correction in the display of figures.

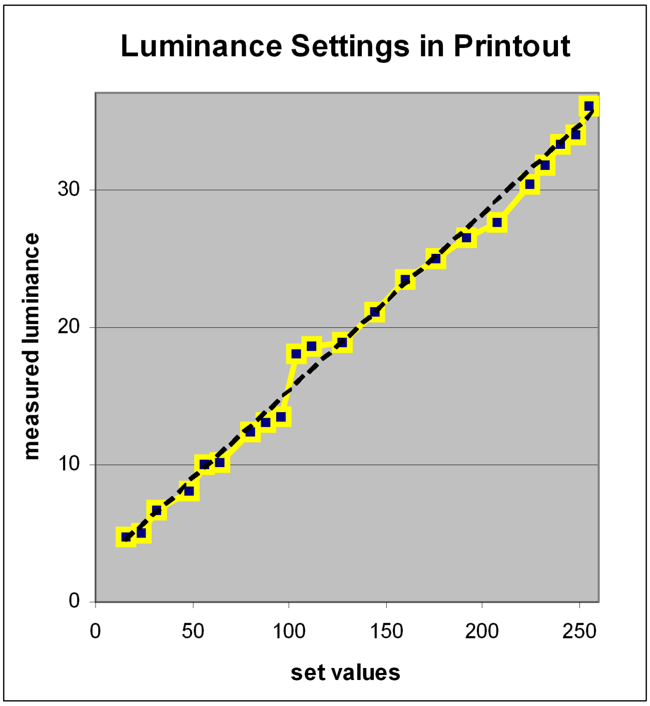


Figure A2. *The "fairly linear" reproduction of a figure like Fig.A1 with a postscript printer*, verified by measuring (under constant illumination) the luminance of printed patches with a photometer. Despite minor deviations, the general course of luminance variations is almost linear–different to the brightness percept of the patches in Fig.A1. If your printer works in a similar way and does not show local contrast enhancement, you may have a good chance to see the "demos" in this paper as intended.

Published online: 6-Jun-2015
For later additions and for comments see http://www.vpl-reports/2/

© *christoph.nothdurft@vpl-goettingen.de*
This document is copyrighted by the author and only for personal, non-commercial use. For any other purposes, please contact the author