The Alejandro Project: Testing the Target Contrast Signal Theory

Replication and Generalisation

Anna Hughes · **Alasdair D. F. Clarke**

Abstract Insert your abstract here. Include keywords, PACS and mathematical subject classification numbers as needed.

1 Introduction

Background Visual search, where participants are asked to find a target within a cluttered scene, has been extensively studied within psychology. A number of models have been developed that can generate testable predictions about how different types of distractors and targets affect search efficiency, including Feature Integration Theory [9] and Guided Search [11, 12]. These models aim to produce a representation of the visual properties of the distractors at each location in the visual field. However, more recent work has taken a different approach, focusing on representing the difference between targets and distractors. For example, in work on eye movement patterns, it has been proposed that performance in inefficient (serial) visual search is mostly determined by the size of the 'functional viewing field', whose size varies as a function of target-distractor similarity [3]. One recent model, the Target Contrast Signal (TCS) Theory [7] aims to provide a unifying, quantitative framework that can make behavioural predictions based on this general assumption.

F. Author
first address
Tel.: +123-45-678910
Fax: +123-45-678910
E-mail: fauthor@example.com

S. Author
second address

Theoretical intro to TCS TCS proposes that behaviour is determined by comparing the target template (held in memory) with every element present in the scene in parallel. This allows the visual system to reject peripheral non-targets quickly; the speed at which items are evaluated is determined by how different the item is from the template through an evidence accumulation process (formally, the slope of the logarithmic function is assumed to be inversely proportional to the overall magnitude of the contrast signal between the target and distractor). The model thus focuses on an initial, efficient processing stage of search; if sufficient evidence is not accumulated during this process, the model posits that a second stage is entered, where attention is deployed serially. TCS has been successful in predicting a number of empirical results, including search performance in heterogeneous scenes based on parameters estimated in homogeneous scenes, both with artificial stimuli [1, 6] and with real-world objects visualised on a computer display [10].

Limitations, and extending basic stimuli While results to date for TCS appear promising, they remain relatively preliminary, being tested by only one research group so far. There remains a great deal of scope for extending beyond the parameters tested to date (such as the number of distractors) in order to test the robustness and generalisability of the model. In addition, there are some predictions of the model, including its ability to explain search asymmetries, that have yet to be empirically confirmed [7].

Within subjects Finally, in all implementations of TCS so far, predictions of search efficiency (e.g. in heterogeneous scenes) have been made on the average of a group of participants, using data from a different group performing a different task (e.g. searching in homogeneous scenes). Thus, we know that TCS can replicate group-level averages between subjects in search well, but we do not know whether it is also able to make predictions at the individual level. This is particularly important given that conclusions based on aggregate data can be different from those that take individual differences into account; in one study where participants searched for a target in an array of randomly oriented line segments, aggregating the data suggested that participants were using a stochastic search model. However, when considering each participant individually, it became clear that there was a high level of heterogeneity in responses, with some participants performing close to optimally, and others actually performing worse than chance [8]. Similarly striking variability has also been reported in other search studies [4, 5].

Conclusion to introduction In the current manuscript, we focus on replicating and extending findings from [2]. Here, participants searched for a target in a scene of homogeneous distractors. First, parallel search efficiency (measured by the logarithmic search slope) was estimated for cases where the distractors varied from the target in one dimension: either colour (e.g. a cyan target being searched for in either yellow, blue or orange distractors) or shape (e.g. a semicircle target in either circle, diamond or triangle distractors). New participants then searched for the same targets in displays where the distractors were compounds, differing from the target in both colour and shape (e.g. searching for a cyan semicircle in either blue circles, orange diamonds
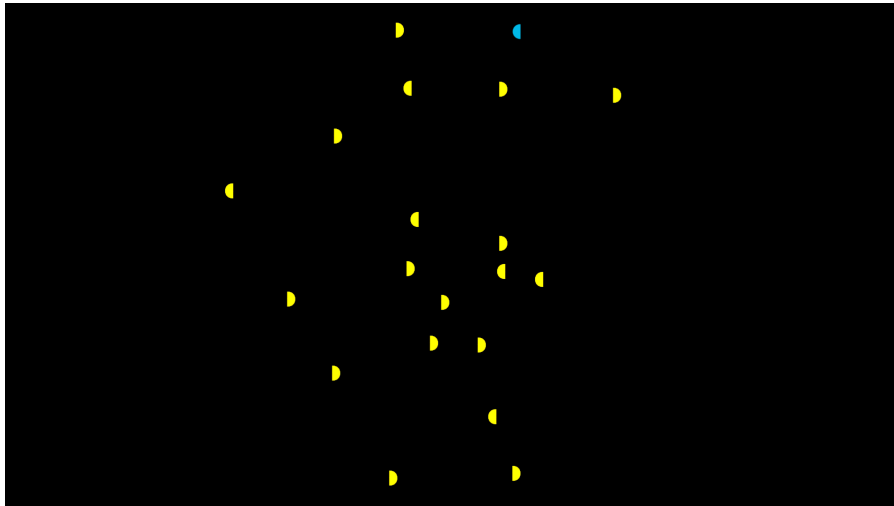
Fig. 1 Example stimulus from [2] Expt 1A. Here, the target is a blue semicircle within a set of homogeneous (yellow semicircle) distractors.

or yellow triangles). Figure 1 shows an example stimulus for the case where the distractors varied from the target in one dimension (from Experiment 1A in their paper). The logarithmic search slopes in the initial experiments were then used to predict the logarithmic slopes and reaction times using a number of models. The authors found that the best model was a 'collinear contrast integration model' where the distinctiveness scores were summed along each attribute in the unidimensional experiments, creating an overall contrast score that was used for compound stimuli predictions.

Overview of the current study We first run a replication of [2], in an online, within-subjects study. This design allows us to extend the modelling, both incorporating a multi-level design to predict within-subjects effects and by utilising a Bayesian generalised linear model framework to better represent the distribution of responses (e.g. avoiding predicting negative reaction times, accounting for uncertainty in model predictions). We will also carry out a direct analytical replication using the same methods as in [2] allowing us to ask whether the choice of analysis affects the results.

2 A multi-level Bayesian reanalysis

Some of the weaknesses of the Target Contrast Signal Theory are i) it doesn't account for individual differences, only the changes to the sample average; ii) it does not account for the distribution of reaction times over multiple trials. Switching from a simple linear regression model to a multi-level model will allow us to compute D for each participant, and estimate the trial-to-trial variance. By also switching from a frequentist to Bayesian framework, we can also account for the uncertainty in the model's predictions. We first describe the original analysis methods, and then our modifications.

2.1 Methods

2.1.1 TCS modelling overview:

According to TCS, Equation 1 allows us to predict mean reaction times:

$$RT = a + \beta \sum_{j=1}^{L} \left((D_j - D_{j-1}) \log \left(\left(N_T - \sum_{i=1}^{j-1} N_i \right) 1_{[2,\infty]}(j) + 1 \right) \right) \tag{1}$$

Where L is the number of distractor types present in the display, N_T is the total number of distractors, N_i is the number of distractors of type i and D_j is the logarithmic slope parameters associated with distractors of type j (from smallest D_1 to largest D_L). The constant a represents the reaction time when no distractors are present and β represents inter-item interactions. When $L = 1$ (all lures are identical), as in the current experiment, the equation simplifies to:

$$RT = a + D_1 \log(N_T + 1) \tag{2}$$

[2] tested three different models for predicting D for compound colour-shape stimuli. The best feature guidance model (Equation 3) suggests that when the target and lures differ in two dimensions, participants will choose to attend to whichever feature dimension is the most discriminable (i.e. has the smallest D value):

$$D_{\text{overall}} = \min\left(D_{\text{colour}}, D_{\text{shape}}\right) \tag{3}$$

The orthogonal contrast combination model instead suggests that independent feature dimensions comprise a multidimensional space, where an object can be described by the overall vector in this space, and thus D_{overall} can be represented as:

$$D_{\text{overall}} = \frac{1}{\sqrt{\frac{1}{(D_{\text{colour}})^2 + (D_{\text{shape}})^2}}} \tag{4}$$

Which simplifies as (we think??):

$$D_{\text{overall}} = (D_{\text{colour}})^2 + (D_{\text{shape}})^2 \tag{5}$$

Finally, the collinear contrast integration model also assumes independence of feature dimensions, but assumes that while the visual features create a multidimensional space, the contrast between them is unidimensional. As D is assumed to be inversely proportional to contrast, the equation can be written as follows:

$$\frac{1}{D_{\text{overall}}} = \frac{1}{D_{\text{colour}}} + \frac{1}{D_{\text{shape}}} \tag{6}$$

[2] found that with their dataset, the collinear contrast integration model was best able to predict their empirical RT results with an $R^2 = 93.27$

We verified we were able to replicate the results of [2] using the dataset available on OSF (https://osf.io/f3m24/)[1] and using the exclusion criteria originally applied; see *Supplementary Materials* for details.

---

[1] downloaded on 28th August 2020
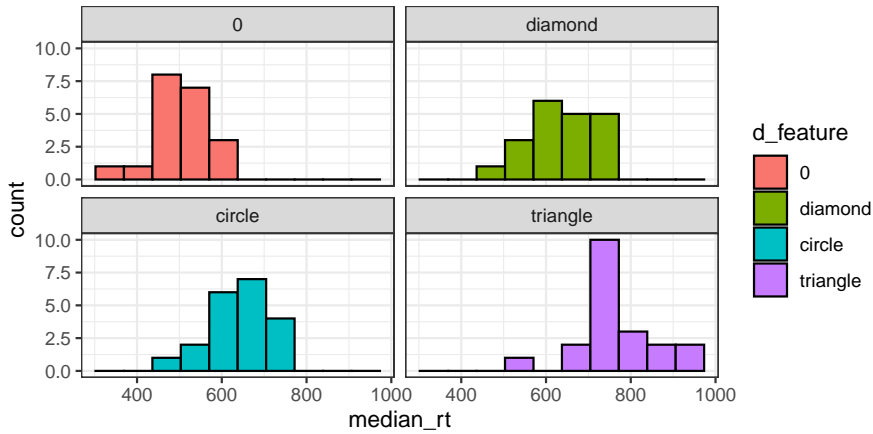
Fig. 2 Histograms of median RTs from Expt 3A in [2]). There are clearly between-subjects differences, as well as differences in distributions for the different conditions.

2.1.2 Our re-analysis approach

What about incorrect trials etc?

First, we computed the D values using a multi-level Bayesian model using a log-normal distribution (using R, brms etc. - see Equation 7). Using this method allows us to fit each individual's performance. This is important because it is clear there are between-subjects differences that are not captured in the original model (see Figure 2).

We fit the model twice, once using a log link[2], and once using the identity link[3]. Other minor changes include modelling reaction times on the seconds scale, rather than milliseconds (making numerical model fitting easier) and removing all RTs less than 10ms.

$$\hat{y} = a + D \log(N_T + 1) \tag{7}$$

The code for our model in *brms* is as follows (possibly this can be a suppmat thing?):

```
m <- brm(
    rt ~  0 + d_feature + log(N_T+1):d_feature +
      (log(N_T+1):d_feature|p_id),
    data = df,
    family = lognormal(link = "identity"),
    prior = my_priors,
    iter = 5000)
```

[2] our preference, as this avoids predicting negative reaction times and helps to account for skew in the distribution

[3] makes this new version of the model more comparable to the original

/0.4/0.4
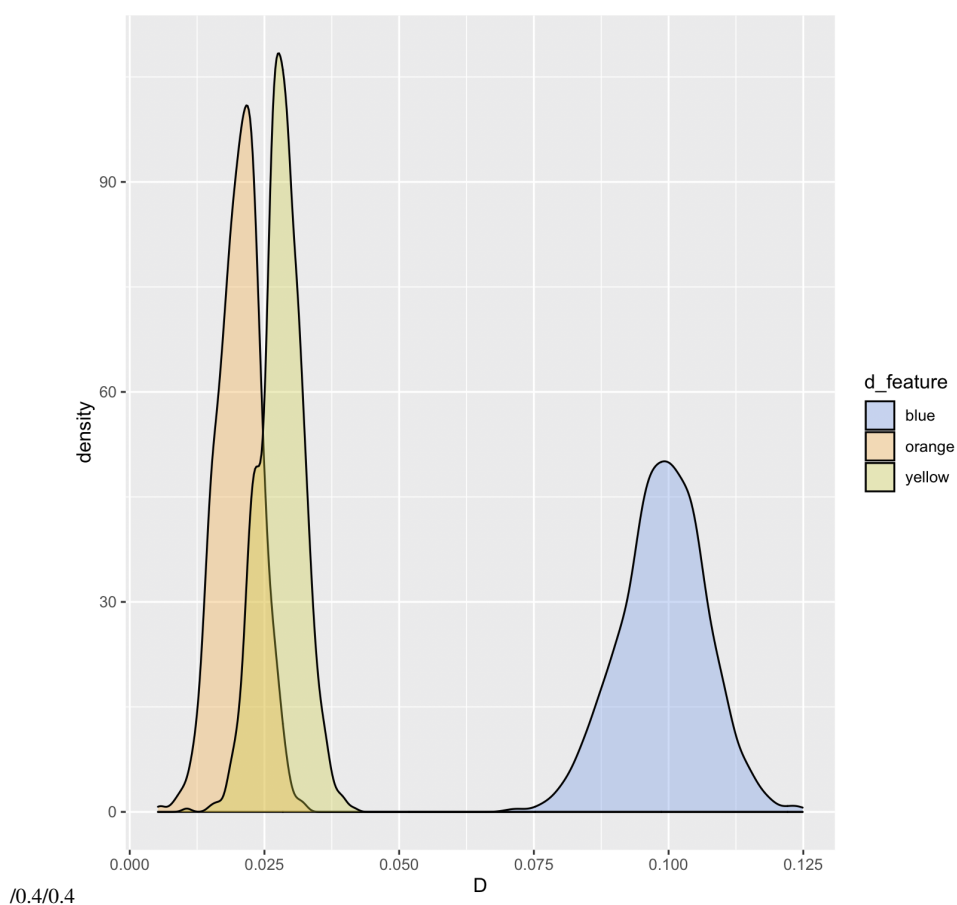
Fig. 3 The shaded regions show the model's predicted D values for each of the conditions in Expt 1A.

2.1.3 Fixed Effects

As we are using a Bayesian framework, rather than D representing the best fit in equation 7, it is now represented as a probability distribution over possible values (see Figure 3). We can can then estimate D in the compound feature distractor cases using the three methods detailed above and compare the predicted values to the empirical values - see Figure 4. Interestingly, the choice of distribution seems to have an effect on which model gives the best prediction; for the normal distribution (as used in [2]) the collinear contrast integration model gives the best fit to the data (as with the original analysis), but for the log-normal distribution, the orthogonal contrast model is better.
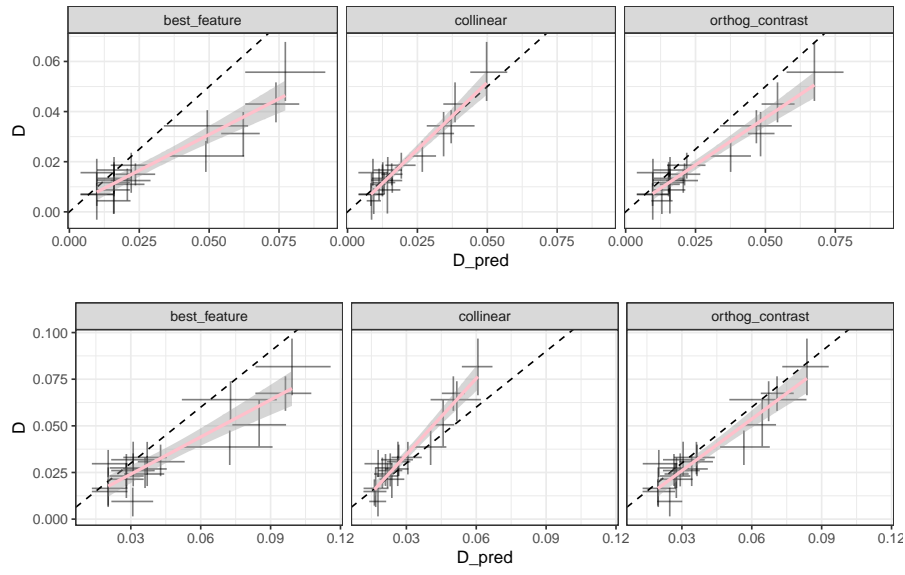
Fig. 4 Correlation between empirical *D* and the predicted *D* using *(top)* normal and *(bottom)* log-normal distributions.
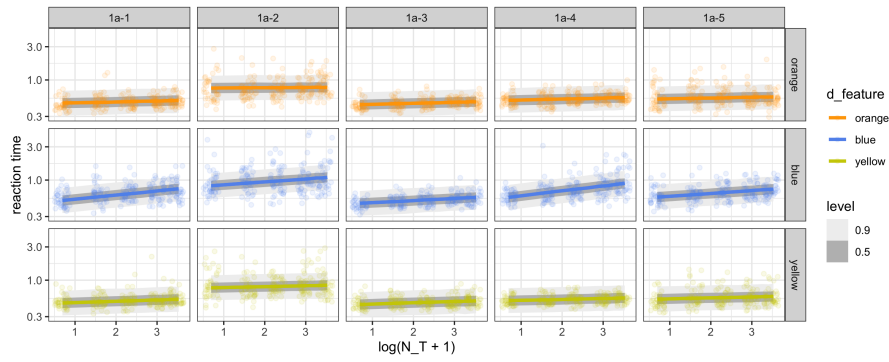
Fig. 5 Example results from the first five participants of Experiment 1a [2]. We can see considerable between-subjects variation.

2.1.4 Individual Differences

While we can't generate within-participant reaction time predictions from the original data, we can present data on the uncertainty around the individual differences in *D* for each condition (see Figure 5).

2.1.5 Power Analysis

To estimate the required sample size for future experiments, we repeatedly reran the above analysis on smaller subsets of the data from [2]. (Full details are given in the *Supplementary Materials*.) Based on this, we found that $n =$ per observer and condition was enough to reliably estimate *D*.

2.2 Discussion

While we replicate the good predictions using the collinear method if we use a normal distribution, we do not find this that this is the best method if we use a log-normal distribution, and instead, we find the orthogonal contrast method is slightly better. This suggests that (seemingly small) analytical decisions may have important implications for the conclusions drawn.

3 Hypotheses

We plan a number of experiments to test the extent to which the original results in [2] replicate and generalise. As well as following the original, between-subjects, experiment design, in which data from one group of observers in one task is used to predict behaviour of a second group of participants in a different task, we will allow for within-subject comparisons. Specifically, we will ask: to what extent do the individual differences in the homogeneous task explain the differences in the heterogeneous task?

1. **Replication of Buetti et al (2019) with online data collection.** Specifically, that the *collinear contrast ingratiation model* outperforms the *best feature guidance*, and *orthogonal contrast combination models*. Furthermore, the $R^2 =$ (99% HPDI $= [,]$) between predicted and observer reaction times.

2. **Larger number of distractors (or targets further in periphery?)** Add an extra ring

3. **Search asymmetries** test O in Qs and Q in Os

4 Experiment 1

H0: confirm collinear method is best (compared to orthogonal contrast combination and best feature guidance models)

[If this is indeed the case, we will use only the collinear contrast integration model in the following experiments].

4.1 Methods

4.1.1 Participants

Based on the power analysis 2.1.5, we will recruit ...

4.1.2 Stimuli

4.1.3 Procedure

4.1.4 Data Pre-processing

incorrect trials? Poorly behaved participants? RTs that are far too short? Or far too long?

4.1.5 Analysis Plan

We will follow the analysis given in Section 2.

4.2 Results

– blank –

4.3 Discussion

– blank –

5 Experiment 2

5.1 Methods

Same as Experiment 1? Perhaps we should have a *General Methods* section for both experiments?

5.2 Results

– blank –

5.3 Discussion

– blank –

6 General Discussion

– blank –

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Simona Buetti, Deborah A Cronin, Anna M Madison, Zhiyuan Wang, and Alejandro Lleras. Towards a better understanding of parallel visual processing in human vision: Evidence for exhaustive analysis of visual information. *Journal of Experimental Psychology: General*, 145(6):672, 2016.
2. Simona Buetti, Jing Xu, and Alejandro Lleras. Predicting how color and shape combine in the human visual system to direct attention. *Scientific reports*, 9(1):1–11, 2019.
3. Johan Hulleman and Christian NL Olivers. On the brink: The demise of the item in visual search moves closer. *Behavioral and Brain Sciences*, 40, 2017.
4. Jessica L Irons and Andrew B Leber. Choosing attentional control settings in a dynamically changing environment. *Attention, Perception, & Psychophysics*, 78(7):2031–2048, 2016.
5. Jessica L Irons and Andrew B Leber. Characterizing individual variation in the strategic use of attentional control. *Journal of Experimental Psychology: Human Perception and Performance*, 44(10):1637, 2018.
6. Alejandro Lleras, Zhiyuan Wang, Anna Madison, and Simona Buetti. Predicting search performance in heterogeneous scenes: Quantifying the impact of homogeneity effects in efficient search. *Collabra: Psychology*, 5(1), 2019.
7. Alejandro Lleras, Zhiyuan Wang, Gavin Jun Peng Ng, Kirk Ballew, Jing Xu, and Simona Buetti. A target contrast signal theory of parallel processing in goal-directed search. *Attention, Perception, & Psychophysics*, pages 1–32, 2020.
8. Anna Nowakowska, Alasdair DF Clarke, and Amelia R Hunt. Human visual search behaviour is far from ideal. *Proceedings of the Royal Society B: Biological Sciences*, 284(1849):20162767, 2017.
9. Anne M Treisman and Garry Gelade. A feature-integration theory of attention. *Cognitive psychology*, 12(1):97–136, 1980.
10. Zhiyuan Wang, Simona Buetti, and Alejandro Lleras. Predicting search performance in heterogeneous visual search scenes with real-world objects. *Collabra: Psychology*, 3(1), 2017.
11. Jeremy M Wolfe. Approaches to visual search: Feature integration theory and guided search. *Oxford handbook of attention*, pages 11–55, 2014.
12. Jeremy M Wolfe, Kyle R Cave, and Susan L Franzel. Guided search: an alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human perception and performance*, 15(3):419, 1989.