Isolating objective and subjective filling-in using the Drift Diffusion Model

Ron Dekel[1,*], Dov Sagi[1,*,#], Ativ Zomet[2], Dennis Levi[3], and Uri Polat[4]

[1]*Department of Brain Sciences, The Weizmann Institute of Science, Rehovot 7610001, Israel.*

[2] *Stanford University, School of Medicine, Pediatrics Hematology Oncology, Palo Alto, CA 94304.*

[3]*Herbert Wertheim School of Optometry & Vision Science, UC Berkeley, Berkeley, CA 94720-2020.*

[4]*School of Optometry and Vision Science, Bar-Ilan University, Ramat Gan 5290002, Israel.*

[*]*Equal contribution*

[#]*Corresponding author. Email: Dov.Sagi@Weizmann.ac.il*

Abstract

Spatial context is known to influence the behavioral sensitivity (*d*') and the decision criterion for detection of low-contrast targets; both were found to interact with reaction time (RT) (Dekel & Sagi, 2020b). We performed an RT analysis of lateral-masking experiments using data from (Polat & Sagi, 2007) and Zomet et al. (2008; 2016). These experiments indicated that the decision criterion, derived from signal detection theory (SDT), increases with increasing target-flanker distance. Our RT analysis revealed that the flanker dependency is stronger with faster RTs and weaker with slower RTs. In Zomet et al. (2008), the observers were slow, because of age or lack of practice; thus, the criterion effects were smaller, showing little or no dependency on RT. These results can be explained by applying the Drift Diffusion Model (DDM), where an evidence accumulation process depends on the flankers via a change in the rate of the evidence, and on priors via a change in the starting point, leading to RT-independent and RT-dependent effects, respectively. The RT-independent distance-dependent bias can be explained by the observers' inability to learn multiple internal distributions (Gorea & Sagi, 2000) required to accommodate the distance-dependent excitatory effects of the flankers, operating in both target present and absent trials.

Introduction

In the present work we analyzed data from previous experiments (Polat & Sagi, 2007; Zomet et al., 2008; Zomet et al., 2016) where observers detected a Gabor target positioned between two high-contrast flankers (lateral masking; (Polat & Sagi, 1993, 1994)). Unlike in the original study (Polat & Sagi, 1993, 1994), where a 2AFC paradigm was used, here only one stimulus frame was presented and observers responded with Yes/No. The latter method allows, under basic SDT assumptions (Green & Swets, 1966), for the computation of a decision criterion. Of relevance is the dependence of the criterion on the target-flanker distance. Motivated by the previously observed interactions between decision criteria in detection experiments with mixed stimuli (Gorea et al., 2005; Gorea & Sagi, 2000, 2001; Zak et al., 2012), we were interested in testing for such interactions in the lateral-masking experiments, where different target-flanker distances were mixed in one block of trials. Following SDT, we consider here two factors affecting the criterion setting: (1) objective, based on the statistical properties of the internal responses to be discriminated (signal and noise likelihood, *Lobj*), and (2) subjective, based on familiarity with the stimuli and task, and on the value of the possible outcome (*Lsbj*). According to Gorea and Sagi (2000), when different targets are mixed in a block of trials, observers cannot independently evaluate the required *Lobj* functions for the different targets. Rather, they use *Lobj* which is based on a mixture of the different distributions involved, even when the different targets are clearly tagged. However, the observers can adjust *Lsbj* when priors are changed, such as when the different targets are presented with different frequencies (prior). In the experiments discussed here, the criterion setting was challenged by mixing stimuli having different target-flanker distances. Indeed, previous results (Polat & Sagi, 2007; Zomet et al., 2008; Zomet et al., 2016) showed that the decision criterion was modulated with distance when distances were mixed, but not so much when the different distances were tested separately.

Our goal here was to identify objective and subjective contributions to this modulation. To this end, we re-examined these results within the context of the Drift Diffusion Model (DDM) (Gold & Shadlen, 2007; Ratcliff & McKoon, 2008; Ratcliff et al., 2016) by incorporating RT measurements that were not analyzed previously. DDM, unlike SDT, introduces a temporal dimension to the decision process, allowing predictions for RT. According to DDM, observers accumulate evidence for and against the presence of a target. Each time interval produces a log-likelihood ratio value $LLR(t)$, assessing the odds ratio for the presence of one stimulus versus the other, which is accumulated across time intervals until a decision is

made when the accumulated value reaches one of two thresholds (bounds), e.g., $+a$ and $-a$, for positive and negative decisions, respectively. The initial value of the accumulator (*sp*, starting point) can be chosen in order to incorporate prior information and the subjective value of the decision (e. g., payoff, reward). The rate of evidence accumulation (drift rate) will increase with the target sensitivity (S/N) and the decision bounds will be reached faster. For a balanced decision process, we assume that the drift-rate (v) is zero when the internal response provides no evidence for target presence or its absence, whereas negative and positive rates point to target-absence and target-presence trials, respectively. Thus, in the context of SDT, the drift rate corresponds to the log likelihood ratio function $LLR(x)=p_S(x)/p_N(x)$, where $p_S(x)$ and $p_N(x)$ represent the distributions of the internal responses (x) in the signal (S) and noise (N) trials.

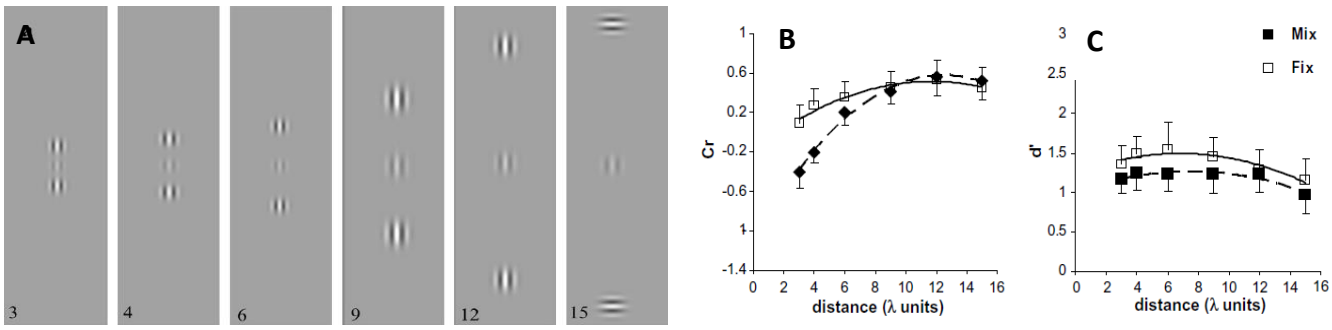

Figure 1: Lateral masking and the SDT criterion. (A) *Stimuli used in Polat & Sagi (2007). Observers detected the presence vs. the absence of the low-contrast target stimulus presented between two high-contrast flankers aligned with the target (except for the 15λ condition used for reference). The distance between the target and flankers was varied (3 ,4, 6, 9, 12, and 15λ) between blocks of trials (the Fixed condition) or within one block of mixed trials (the Mixed condition).* ***(B and C)*** *Group results from Polat & Sagi (2007, Fig. 3): the effect of flankers on target detection, as measured by the signal detection theory (SDT)* ***(B)*** *criterion (Cr), and* ***(C)*** *sensitivity (d'). Shown are the means ± SEM across 7 observers (see the Methods). Note the close to uniform criterion level in the Fix condition compared with the larger range in the Mixed condition. d' does not differ much between conditions, though it is somewhat lower in the Mix condition, possibly due to the increased uncertainty involved in this condition. The d' curves show lateral facilitation at shorter distances, though smaller than the typical facilitation observed with the standard 2AFC method (Polat & Sagi, 1993; Polat & Sagi, 2007).*

Interacting decision criteria

Consider the case where different targets are mixed during an experiment, producing different drift rates, possibly requiring independent accumulators for efficient detection. The findings of Gorea and Sagi (2000) suggest that only one accumulator is used in the mixed condition, with the estimated evidence for/against target presence being blind to the originating distribution. Thus, we expect zero evidence, and the implied decision criterion, to correspond to the presence/absence of targets in any of the possible flanker configurations, i.e., to be globally set based on a mixture of the different distributions (related to the different target-flanker distances; see Figure 2).

In the present work we analyzed the task of detecting a low-contrast Gabor patch in the presence of flankers (Fig. 1A). The effect of the flankers depends on the target-flanker distance, with improved sensitivity at intermediate target-flanker distances ($2\lambda < \text{distance} < 8\lambda$) (where λ is the wavelength of the Gabor stimuli). This range-dependent lateral facilitation (Polat & Sagi, 1993), between co-aligned Gabor stimuli (Polat & Sagi, 1994), produces a subjective impression of filling-in, leading to a range-dependent increase in the False Alarm (FA) rate in Yes/No detection experiments (Polat & Sagi, 2007). The subjective effect can be quantified by the decision criterion (Fig. 1B) as defined in signal detection theory (SDT) (Green & Swets, 1966). Although the criterion can be adjusted by observers when different target-flanker distances are blocked (Polat & Sagi, 2007), it was found to depend on the flankers when the different target-flanker distances are fixed (Fig. 1C) (Amiaz et al., 2011; Polat & Sagi, 2007; Zomet et al., 2008; Zomet et al., 2016). This result can be explained by observers failing to independently adjust multiple criteria, as illustrated in Figure 2. To explain these effects, the flankers can be assumed to affect the gain and the offset of the target stimulus population, with the effect depending on the target-flanker distance. This immediately leads to the effect on the criterion illustrated in Figure 1B (Polat & Sagi, 2007), and it also explains the improved sensitivity due to the non-linearity of the stimulus vs. the internal response function at low contrasts (Chen & Tyler, 2001). Figure 2 illustrates the situation faced by observers in such a task. Each trial leads to an internal response drawn from one of 12 distributions, corresponding to the 6 distances with and without the target present (6x2, see Fig. 2A). For an observer to make unbiased decisions, these twelve distributions need to be correctly estimated, with an unbiased criterion set for each distance based on the corresponding pair of distributions (Noise and Signal). Instead, we suggest that observers utilize only two distributions, the average Noise distribution and the average Signal distribution (the black curves in Fig. 2A), and that their ratio will provide the input to their decision

(Yes if LLR(x)>1, No otherwise, ‘x’ being the internal response, see the black curve in Fig. 2B). For a blocked condition (a Fixed distance), observers can estimate the specific distance-dependent distribution and derive an unbiased likelihood ratio value for each distance (the colored curves in Fig. 2B). Figure 2C depicts the predicted decision criteria for the example presented in Figure 2A.

It is also possible that observers shift their criteria for each distance based on prior knowledge about the statistics of filling-in between contour elements, as shown by Gorea and Sagi (2000) for targets having different prior probabilities. Such shifts may affect the predictions shown in Figure 2C, depending on the specific priors assigned to different distances. Thus, as described above, we can expect two sources to contribute to the criterion effect: (i) an objective one based on the statistics of internal responses, and (ii) a subjective one based on observers’ prior knowledge concerning the stimuli and the task at hand.

A criterion that disappears with RT and that does not disappear with RT

We have previously shown for some tasks (detection, Tilt After Effect, and Tilt Illusion) that perceptual biases due to shifts in decision criterion vanish at longer RTs, whereas biases due to sensitivity changes are probably not affected by RT (Dekel & Sagi, 2020b). Such effects can be explained by assuming that the brain performs these tasks using a gradual accumulation of noisy evidence, as in the standard drift diffusion model (DDM) (Gold & Shadlen, 2007; Ratcliff & McKoon, 2008; Ratcliff et al., 2016). In these models, shifts in decision criterion due to priors and payoffs (subjective) are implemented by a change in the starting point of the evidence accumulation process; therefore, when the process takes longer to terminate (a slower RT), decisions are less biased due to noise accumulation (Dekel & Sagi, 2020b). However, changes in decision criteria due to sensitivity changes are modeled as a change in the rate at which evidence is accumulated (integrated likelihood); thus, such effects are largely RT independent. Here, we apply this approach to the lateral masking data collected for existing publications but previously not analyzed for RT (Polat & Sagi, 2007; Zomet et al., 2008; Zomet et al., 2016). The experimental results and the unique distribution model show a clear dependency of decision bias on the target-flanker distance, a dependency that is expected to disappear at slower response times if caused by subjective factors such as learned priors, but not if caused by a failure to form efficient representations of the internal distributions associated with the different stimuli.

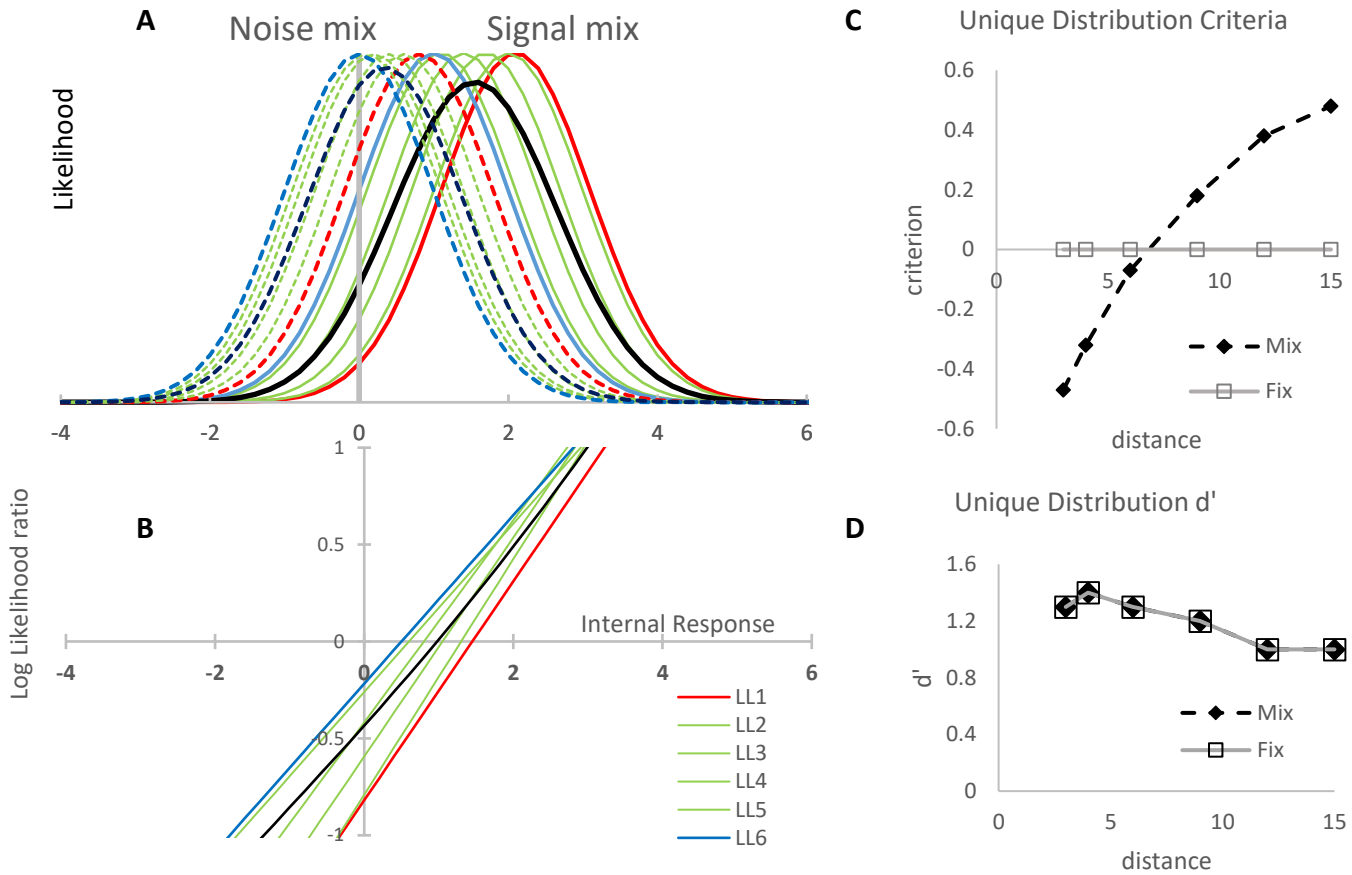

Figure 2: The proposed SDT modeling of the measured criterion change. *(A) Internal response distributions thought to be involved in the studied task (normal distributions, σ=1). There are 6 noise distributions (broken lines) and 6 signal distributions (continuous lines), with the shortest and longest target-flanker distances denoted in red and blue, respectively. We assume that the observers cannot estimate all of these distributions when encountered randomly within a block of trials, and instead use the mean distributions of Noise and Signal (the broken and continuous black curves, respectively). (B) Log likelihood ratio (LLR) functions computed for Noise and Signal pairs in A. We expect a 'Yes' response when LLR>0 and 'No' otherwise. Observers are assumed to base their decisions on the black curve (LLR of the means) instead of the specific distance curves; thus, they are biased. (C) Predicted criterion shifts derived from B. For the Fix condition, we assume ideal observers (unbiased). (D) d' values associated with the distribution pairs are presented in A. These values are computed as the difference between the paired Signal and Noise distributions.*

Methods

Experimental data

Here we analyzed unpublished RT data from previously reported experiments (Polat & Sagi, 2007; Zomet et al., 2008; Zomet et al., 2016), as summarized in Table 1. All experiments measure the detection of low-contrast Gabor patches in the presence of lateral flankers (Fig. 1A). In all experiments, the 15λ distance used horizontally oriented flankers, presumably nulling any lateral interaction with the vertical target. From Polat and Sagi (2007), we managed to recover the original data for most observers (six and five observers out of seven, for Mix and Fix, respectively). In the remaining publications, all the original data were recovered, as well as an unpublished pilot study with numerous repetitions per participant. The observers in all experiments differed, with the exception of those in Polat and Sagi (2007), where five observers were shared between Fix and Mix. The stimulus parameters were nearly identical in the different experiments (see the differences in Table 1).

Table 1 – Summary of the experimental data.

Publication	Experiment	Design				Participants			
		Target-flanker distance (λ)	External noise levels	Mixing of target-flanker distances	Negative feedback	N	Gender (M/F)	Age (Mean ± *SD*)	#Trials per participant (Mean ± *SD*)
Polat & Sagi (1993)	Mix	3,4,6,9,12,15*	-	Mixed	Y	6**	5/1	17 ± 2	3550 ± 1400
	Fix	3,4,6,9,12,15*	-	Blocked	Y	5**	4/1	17 ± 2	5400 ± 2800
Zomet et al. (2008)	Controls	3,4,6,9,12,15*	-	Mixed	Y	32	14/18	42 ± 15	110 ± 130
	Depress	3,4,6,9,12,15*	-	Mixed	Y	27	15/12	48 ± 15	120 ± 190
Zomet et al. (2016)	Main experiment	3,4,6,15*	0,0.5,1,2	Mixed	Y	12	2/10	24 ± 3	930 ± 50
	Unpublished pilot	3,4,6,9,12,15*	0,0.5,0.75, 1,1.5,2	Mixed	Y	7	2/5	26 ± 4	2100 ± 50

* The 15λ distance uses flankers with a horizontal orientation.
** Missing data from the original study.

Signal detection theory (SDT) analysis

We used the standard definitions of the sensitivity (*d*') and the internal criterion (*c*) (Green & Swets, 1966):

$$d' = z(\mathrm{P_{Hit}}) - z(\mathrm{P_{FA}}) \tag{1}$$

$$c = -0.5 \cdot [z(\mathrm{P_{Hit}}) + z(\mathrm{P_{FA}})] \ , \tag{2}$$

where $\mathrm{P_{Hit}}$ is the probability that an observer correctly reported that the target is present in target-present trials, $\mathrm{P_{FA}}$ is the probability that the observer incorrectly reported that the target is present in target-absent trials, and z is the inverse cumulative normal distribution function. To avoid saturation, the $\mathrm{P_{Hit}}$ and $\mathrm{P_{FA}}$ probabilities were clipped to the range $\left[\frac{1}{2n}, \frac{2n-1}{2n}\right]$, where n is the number of trials in the measurement.

Reaction time (RT) behavioral analysis

Trials were binned by RT into approximately equal-quantity bins. Binning was done separately for each combination of observer × experimental block × target-flanker distance × target stimulus (present/absent). Using such a strict binning rule prevents confounds due to interactions between stimuli or the condition and RT.

Statistics

All statistics were assessed using linear mixed-effects models, as detailed in the text. All models had an additive individual component (offset), conceptually similar to an ANOVA that has a repeated-measures individual factor.

Results

We binned trials by categorizing RT into four equal-quantity bins from the fastest to the slowest. Binning was done separately for each experimental block and each trial type to avoid confounds (see the Methods). Task performance was quantified in each RT bin using the standard measures of sensitivity (*d*', Eq. 1) and decision criterion (*c*, Eq. 2) from signal detection theory (SDT) (Green & Swets, 1966).

We will first consider the experimental data of the "Mix" condition from Polat and Sagi (2007). In this experiment, trials having different target-flanker distances were mixed within a block (Fig. 1A). In this design, the original results showed flanker-dependent criterion modulation, as seen in Figure 1B (see the Methods). Importantly, the analysis here revealed a tight link between RT and criterion. As shown in Figure 3A, the criterion (*c*) had lower values for slower RTs, especially at larger target-flanker distances (λ), with all distances having $\lambda \geq 6$ showing $p < 0.02$ with $t_{(22)} < -3.36$ for an effect of RT bin index on *c* (using linear mixed-effects regression, Bonferroni corrected for six multiple comparisons). We interpreted these observations to indicate two effects: First, the overall criterion (ignoring the target-flanker distance) is reduced with increasing RT ($t_{(142)} = -5.08$, $p = 1\times10^{-6}$). Second, the criterion slope as a function of distance is reduced with increasing RT ($t_{(26)} = -5.00$, $p = 3\times10^{-5}$). (A different statistical approach for the second claim is that the interaction between the RT bin index and distance is significant: $t_{(140)} = -2.96$, $p = 4\times10^{-3}$.) Importantly, despite the reduction, the slopes remained positive in all RTs, even the slowest (all RT bins showed $t_{(34)} = 2.85$ with $p < 0.03$ for a linear modulation of *c* by λ, Bonferroni corrected for four multiple comparisons).

Polat and Sagi (2007) found a small but significant modulation of sensitivity (*d*') by the flankers, as seen in Figure 1C for the available data. (Note that a much stronger modulation of sensitivity by the flankers is found when the detection task is replaced by a 2AFC discrimination task (Polat & Sagi, 1993, 2007).) Applying the RT analysis here revealed a gradual reduction of *d*' in slower RTs, with values decreasing from ~1.5 in fast RTs, to ~1 in slow RTs (Fig. 3C; $t_{(142)} = -4.01$, $p = 1\times10^{-4}$ for modulation of *d*' by RT when ignoring the target-flanker distance). There was no significant interaction between RT and the target-flanker distance ($t_{(140)} = -0.72$, $p = 0.5$). Note that criterion and sensitivity exhibited different modulations by RT: the criterion was modulated to a much greater extent than sensitivity, and its modulation dynamics differed (Fig. 3A vs. Fig. 3C).

In an additional experiment by Polat and Sagi (2007), the trials were blocked by the target-flanker distance (the "Fix" experiment). Here, observers were expected to better adjust their criterion for the different

decision requirements presumably imposed by the different target-flanker distances. The results of the RT analysis (Fig. 3) were similar to the "Mix" experiment. Specifically, criterion modulation by RT (Fig. 3B) was significant for $\lambda \geq 9$ ($p < 0.01$ corrected), with a significant overall reduction ($t_{(118)} = -5.63$, $p = 1\times10^{-7}$), and a significant reduction with RT in the criterion slope as a function of the target-flanker distance ($t_{(18)} = -4.42$, $p = 3\times10^{-4}$). The modulation of criterion by the flanker distance (slope) was evident in all RT bins, though it was only statistically significant in faster RTs (two faster RT bins: $t_{(28)} \geq 3.22$, $p \leq 0.01$ corrected; two slower RT bins: $t_{(28)} \geq 1.30$, not significant). Sensitivity (Fig. 3D) showed an overall reduction with RT ($t_{(118)} = -2.84$, $p = 5\times10^{-3}$), but no significant interaction between RT and target-flanker distance ($t_{(116)} = -0.56$, $p = 0.06$).

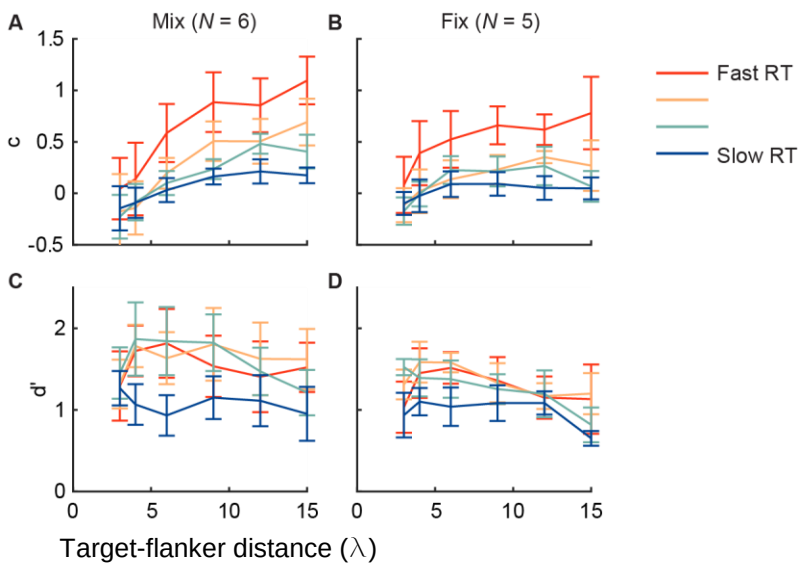

Figure 3: Target-flanker distance and RT. *Analysis of (Polat & Sagi, 2007).* ***(A and C)*** *Shown are* ***(A)*** *criterion (c) and* ***(C)*** *sensitivity (d'), as a function of the target-flanker distance (in units of wavelength, λ), when the different target-flanker distances are mixed within blocks ("Mix" experiment). Data were split into four bins (colors) by sorting the measured RTs. The results indicated: (i) a more biased criterion setting in the faster RTs, (ii) a weaker modulation of the criterion by the flankers in the slower RTs, and (iii) a flanker-dependent criterion in all RTs, even the slowest. The sensitivity (d') was lower with slower RTs, but only in the slowest time bin (red).* ***(B and D)*** *Behavioral data of the "Fix" experiment, whereby trials are blocked by the target-flanker distance, showing reduced effects. Error bars are ±1 SEM.*

The external noise abolished the RT effects, even when sensitivity was fixed

Next, we consider experiments whereby different levels of external noise are added to the target (see the Methods, in Zomet et al. (2016)) using a procedure that is otherwise identical to the one considered in the previous section. Different noise levels were added in different blocks, randomly permuted between observers (see the Methods, Zomet et al. (2016)). We consider two variants: one where contrast was maintained when noise was increased, leading to reduced *d*' when the noise level increased ("Main", N = 12 observers, Fig. A5). The other, an unpublished pilot of that study, whereby the target contrast was increased with noise to maintain a roughly fixed *d*' ("Pilot", N = 7, Fig. A6; this pilot is particularly convenient for the purpose of the RT analyses performed here by virtue of having a large number of trials per observer, see Table 1).

First, we consider the case when the level of external noise was zero ("No noise", Figs. 4BC, A3, A4). In this condition, the experiment is equivalent to the one considered in the previous section (Mix condition), so we expect similar results. The results indeed show a similar interaction between criterion and RT. Specifically, the overall criterion was reduced with RT ("Main": $t_{(164)} = -4.71$, $p = 5\times10^{-6}$; "Pilot": $t_{(188)} = -3.38$, $p = 9\times10^{-4}$), and the criterion slope as a function of distance was also reduced with slower RTs ("Main": $t_{(46)} = -4.19$, $p = 1\times10^{-4}$; "Pilot": $t_{(26)} = -5.00$, $p = 3\times10^{-5}$).

Next, we consider the influence of the added noise. Noise was added in multiples of the individual threshold (between 0 and 2 times the threshold) (Fig. 4A). As shown in Figure 4B, introducing noise abolishes the interaction of criterion and RT. Moreover, this effect is not due to a reduction in sensitivity (*d*'), as shown in Figure 4C, where the target contrast was increased with the added noise, so sensitivity was fixed. Interestingly, at all noise levels, it can be observed that the influence of the flankers remains approximately the same. Indeed, in the slowest RT bin, the differences between experiments and conditions are largely abolished, as clearly seen in Figure 5 when comparing bin 4 (a slow RT) with bin 1 (a fast RT).

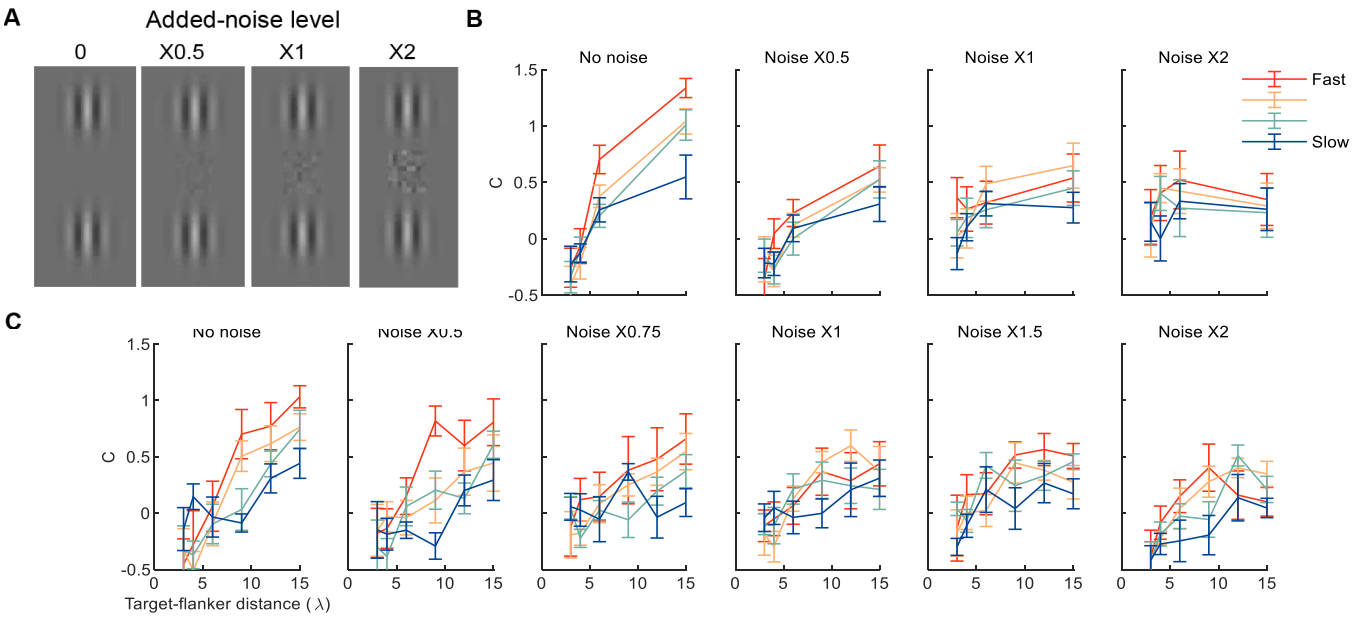

Figure 4: External noise and the RT effect. *Analysis of Zomet et al. (2016). (**A**) Example stimuli with different added noise levels. (**B**) The "main" experiment where added noise reduces sensitivity (d'). (**C**) The "pilot" experiment where target contrast is increased with noise, so that sensitivity remains stable across noise levels. The error bars are ±1 SEM.*

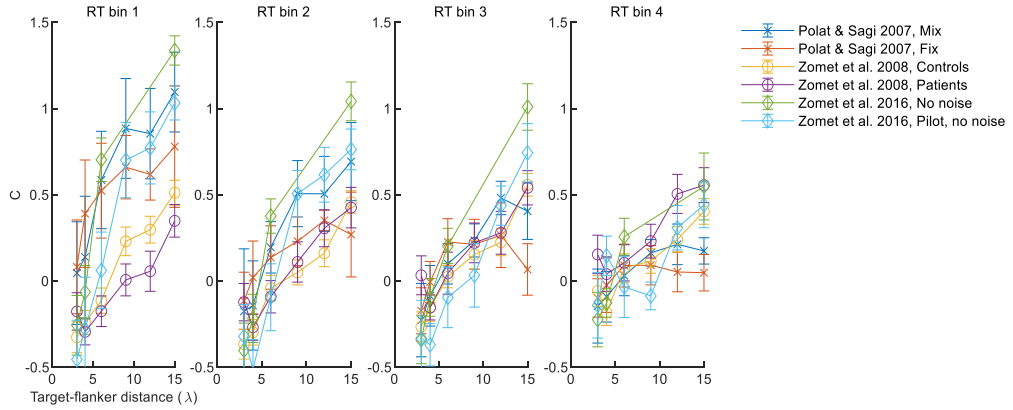

Figure 5: The differences between experiments disappear with slow RTs. *Shown for all experiments are the criterion measurements as a function of the target-flanker distance in four RT bins. It can be observed that the experimental differences are abolished in the slowest RT bin (bin 4).*

Slow individuals show no RT effect for criterion

Next, we applied the RT analysis to the experimental data of Zomet et al. (2008). In this study, patients diagnosed with depression ("Patients") and their matched controls ("Controls") performed an experiment similar to the ones analyzed above. Specifically, different target-flanker distances were mixed within a block, and there was no noise, comparable to the "Mix" experiment of Polat and Sagi (2007) and the no-noise condition of Zomet et al. (2016). Unlike these studies, the number of trials was lower, and observers were, on average, older (see Table 1).

The results of the RT analysis (Fig. 6A) showed an overall criterion that was not significantly modulated by RT ("Controls": $t_{(766)} = 1.06$, $p = 0.3$), or even significantly *increased* with RT ("Patients": $t_{(646)} = 5.61$, $p = 3\times10^{-8}$). This is unlike the findings of the experiment having a similar design that showed a reduction in criterion with RT (Figs. 3 and 4). The other effects were more in line with previous findings. The criterion slope as a function of the target-flanker distance was reduced with RT, though not appreciably (Fig. 6A; "Controls": $t_{(126)} = -3.06$, $p = 3\times10^{-3}$; "Patients": $t_{(106)} = -0.62$, $p = 0.5$). The sensitivity was strongly reduced with RT ("Controls": $t_{(766)} = -5.73$, $p = 1\times10^{-8}$; "Patients": $t_{(646)} = -10.94$, $p = 1\times10^{-25}$), independently of the target-flanker distance ("Controls": $t_{(764)} = -0.20$, $p = 0.8$; "Patients": $t_{(644)} = -1.36$, $p = 0.2$).

To explain the small effect of RT on criterion, we considered the simple idea described in Dekel and Sagi (2020a): that individuals with reduced bias can be explained by having slower decision times. Indeed, we found that observers in Zomet et al. (2008) had extremely slow RTs, showing, on average, ~1600 ms for Controls, and ~2700 ms for Patients (see the x-axis of Fig. 6B). This is much slower than the ~650 ms measured for observers in similar experimental conditions (Fig. 3A and Fig. 4 "No noise"). Overall, the differences between experiments can possibly be explained by RT. In addition, we considered differences between observers in the "Controls" group of Zomet et al. (2008), where wide inter-individual differences in RT were measured (see the x-axis of Fig. 6B). We correlated individual RT with the individual size of criterion modulation by RT (for 15λ distance), finding a strong correlation (adjusted r=0.54, p=5x10^{-8}; Fig. 6B). To summarize, the difference between the experiment by Zomet et al. (2008) and the other experiments appears to be well explained by slower RTs. In addition, the difference in RT between experiments and observers can possibly be attributed to age and/or practice effects (Table 1).

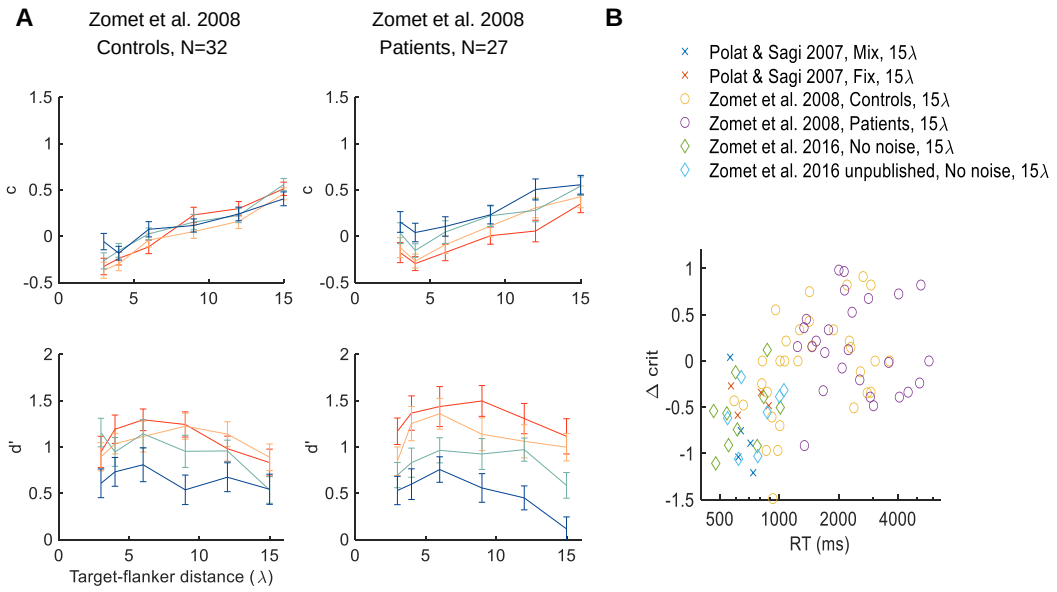

Figure 6: The RT effect in slow individuals. (A) *RT analysis of Zomet et al. (2008). Annotations follow Figure 3. The results showed no RT effect for criterion in depressive patients and controls. (**B**) Shown for each observer is the change in criterion (Δc) from the fastest to the slowest RT bin of the 15λ measurement, as a function of the mean RT. The results indicated a correlation between RT and Δc between experiments, and within experiments for the* Zomet et al. (2008) *control data. The procedural differences between experiments are summarized in Table 1.*

Summary of results

The experimental results show a clear dependency of decision bias on the target-flanker distance, which is expected to disappear at slower response times if caused only by subjective factors such as learned priors, but not if caused by a failure to form efficient representations of the internal distributions associated with the different stimuli. Our results, summarized in Figure 7, support the latter.

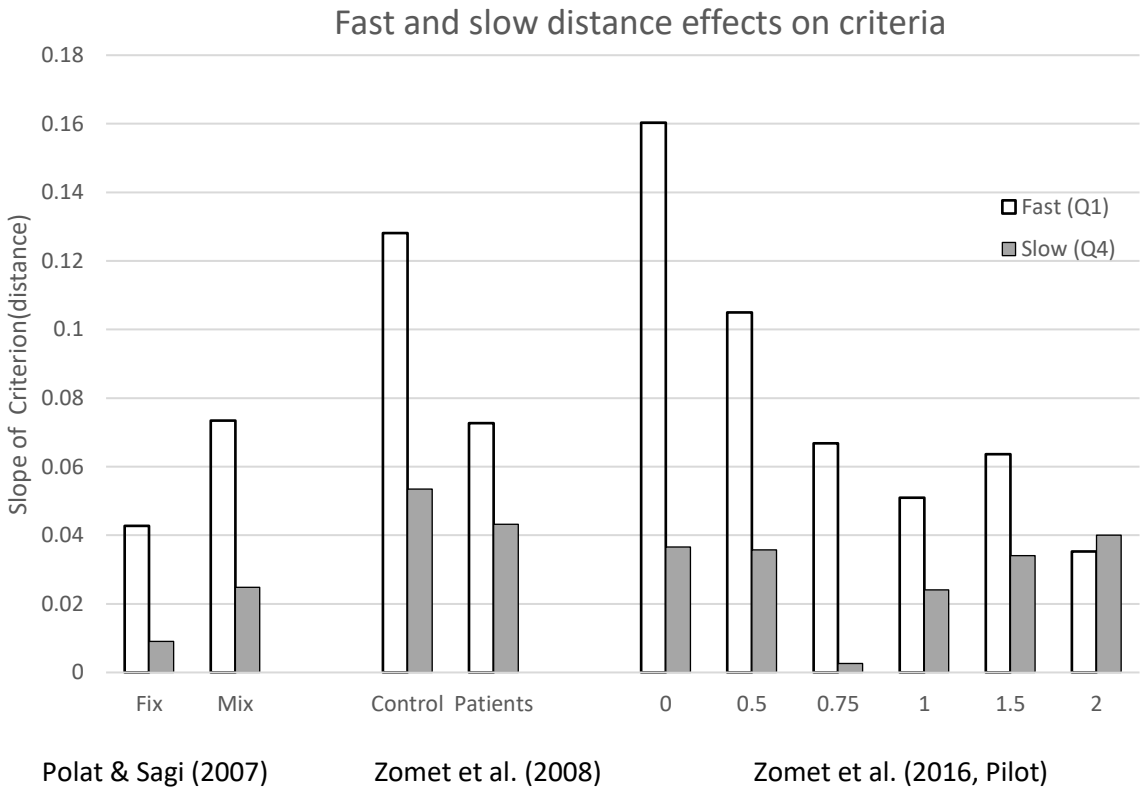

Figure 7*: Fast (Q1: 1st RT quantile) and slow (Q4: 4th RT quantile) distance effects on criteria: a summary of three studies. Clearly, the fast criterion effects (empty bars) are larger than the slow criterion effects (filled bars). All fast distance effects are highly significant (see the Results and Figures A5 and A6). The slow effects are more uniform across conditions, except for the Fix condition (Polat & Sagi 2007, though statistically significant, see the Results) and the noise amp=0.75 condition (Zomet et al., 2016, Pilot, not significant). RT effects (Q1 vs Q4) do not reach statistical significance in the patient's group and in noise experiments with noise amplitude > 0.5; otherwise, they are highly significant (see the Results and Figures A5 and A6).*

DDM predictions

To explain the observed interactions with RT, we considered the idea that the detection task is performed by gradual accumulation of noisy evidence, as in the standard drift diffusion model (DDM) (Gold & Shadlen, 2001; Ratcliff et al., 2016). Here we considered some general effects of model parameters on its behavior (more detailed model parametrization and fitting results are provided in the Appendix, Fig. A2). Two important parameters in the model affect the criterion: (i) the starting point of the accumulation process, which leads to less bias in slower decisions (Fig. 8, second row) (Dekel & Sagi, 2020b) and (ii) the rates at which evidence is accumulated for target-present and target-absent trials (Fig. 8, first row). As shown in Figure 8, this simple idea can explain the criterion measurements quite well. We provide here some intuition as to how this model works by describing the predictions of four simple model variations and analyzing the effect of the drift rate and the starting point on the results.

<u>Model 1: An unbiased flanker-independent starting point, target-dependent drift rate.</u> Here, the drift rate (v) depends on the target stimulus (v_+=+1 for target present, v_-=−1 for target absent), and the starting point (sp) is fixed at 0.5 (unbiased). This model predicts no bias (c=0) at all RTs, with fixed d', independent of distance. This model fails to explain the data.

<u>Model 2: As in Model 1 with a flanker-dependent drift rate.</u> Here, at short distances, the difference between v_+ and v_- is increased, thus showing larger d' values at these distances. Both v_+ and v_- are positive at short distances (supporting positive responses), thus showing a negative bias (c<0) at these distances, RT independent. This model can explain the distance dependence of the criterion but not the RT effect.

<u>Model 3: A biased flanker-independent starting point (sp), flanker-dependent drift rate.</u> The same as in Model 2, but with a biased starting point (sp). The effect of this bias (sp<0.5) is short term, resulting in higher criteria (c) in trials with fast RTs, and with slow RTs showing criteria equal to those produced by Model 2. The RT effect on c is distance independent. This model predicts an RT effect but cannot explain the dependence of the criterion slope on RT (Fig. 7).

<u>Model 4: A flanker-dependent starting point (sp) and drift rate.</u> Like Model 3 but with sp allowed to change with distance, resulting in distance-dependent RT effects (RT-dependent slope of the *c(distance)* curve). This model captures the main features of the experimental results.

Overall, the DDM framework considered here accounts for the most salient RT effects, although the precise modeling details remain somewhat speculative (see the Discussion, and Fig. A2).

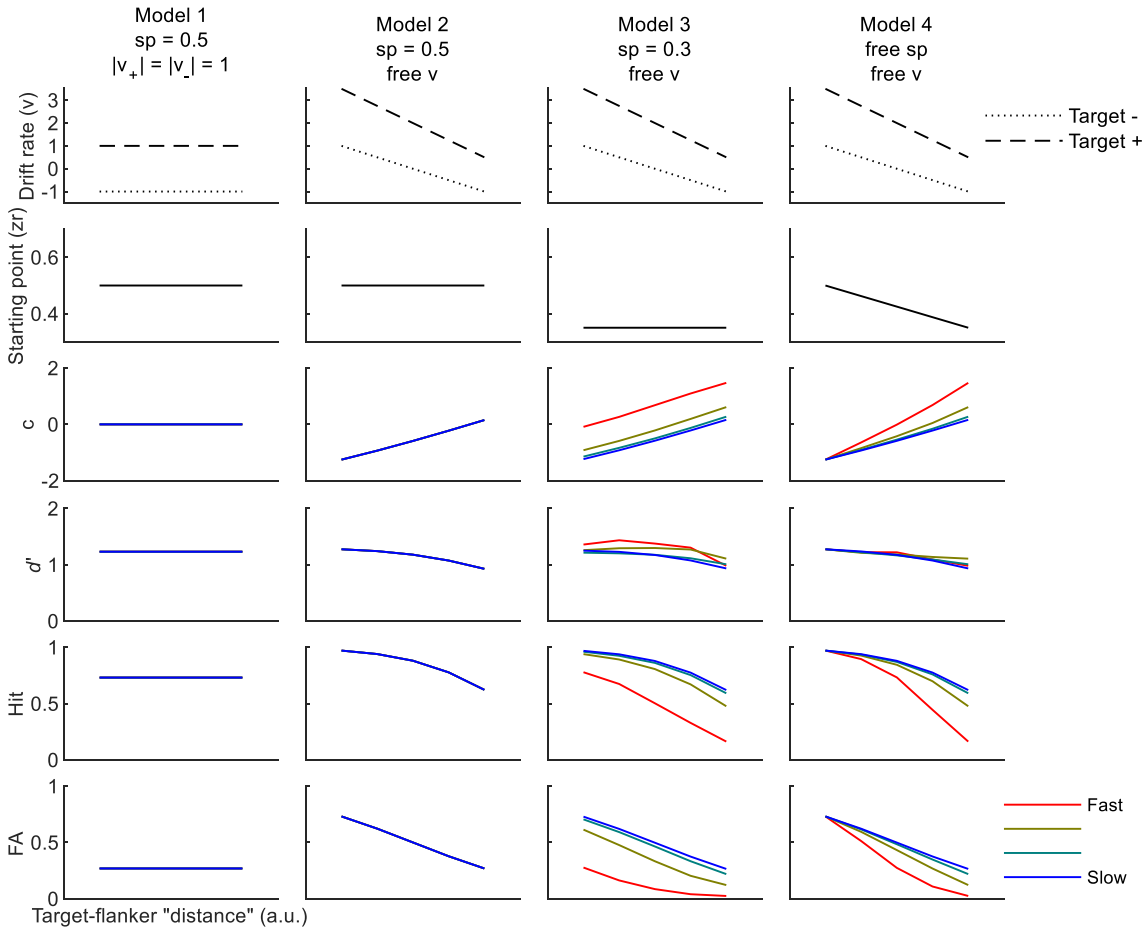

Figure 8: Modeling the influence of flankers on target detection using DDM. *Shown are the parameters (the first two rows) and the theoretical predictions (the last three rows) of the four models described in the text. The sensitivity estimate (d') is approximately the difference in drift rate between the target present and absent trials. In models 3 and 4, the starting point is biased (sp < 0.5), leading to an overall reduction in criterion (c) with RT. In Model 4, the reduction in slope with RT can be explained by a flanker-dependent starting point (variable sp). Overall, Model 4 manages to explain both the reduction in overall criterion with RT, and the reduction in slope with RT. See the text for model details and Fig. A2 for fitting Model 4 to data.*

DDM drift rate and SDT *d*'

Note that the DDM analysis allows for the derivation of target-flanker distance-dependent sensitivity functions for both target-present and target-absent conditions (the idea in Fig. 8, fitting in Fig. A2), unlike the previously used SDT methods that provided only a differential sensitivity estimate (i.e., *d*'). As seen in Fig. A2, the fitted drift rates for both target-present and target-absent both gradually decrease with longer target-flanker distances, as predicted by the SDT theory illustrated in Figure 2. The differential sensitivity, which is the difference in the fitted drift rate between target-present and target-absent trials, was mostly fixed, showing a small decrease at longer target-flanker distances. This is consistent with the slightly improved *d*' in the presence of flankers. We noted that the fitted drift rate in the target-absent trials was usually negative. A negative drift indicates that at each time point, in target-absent trials, the evidence supports the negative (target-absence) response. That is, the reduced sensory response in the target-absence trials is mapped into *negative*, not zero, which favors target-presence (see the Discussion). This finding is reasonable in the sense that what is accumulated is target-presence evidence, with "zero" values corresponding to no evidence for or against target presence (Ratcliff & Smith, 2004).

Discussion

Lateral masking affects the detection of low-contrast targets (Polat & Sagi, 1993). This effect is usually quantified using measures of behavior derived from signal detection theory (SDT) (Green & Swets, 1966): sensitivity (*d*', Eq. (1)) and decision criterion (*c*, Eq. (2)). Although the bias-independent sensitivity measure (*d*') is the standard measure used to quantify perceptual effects, our interest here lies in effects on decision criterion (bias), shown in Figure 1. SDT, with limited access to internal distributions (Gorea & Sagi, 2000), provides an observer-independent account of the biases found, as illustrated in Figure 2 (see the Introduction). Here, for the first time, we performed an RT analysis of the lateral masking effect, attempting to isolate the subjective (observer depended) and objective (observer independent) factors underlying the observed biases in perception. To model RT, we used the drift diffusion model (DDM), which allows for the derivation of internal response parameters in trials with and without target presence. The recovered parameters are mapped to objective and subjective factors affecting perception in the presence of spatial context.

We used data from Polat and Sagi (2007) and Zomet et al. (2008; 2016). These studies show that the decision criterion increases with target-flanker distance (for relevant ranges $>2\lambda$). The results of the RT analysis here showed that this dependency is stronger with faster RTs and weaker with slower RTs. In addition, the overall criterion became more unbiased with slower RTs, and, consequently, differences between experiments regarding the criterion setting largely disappeared with slower RTs. These findings suggest a stereotypical flanker-dependent criterion modulation that is revealed by the slow RTs. The target sensitivity showed modulation by the flankers, as well as deterioration in slower RTs, having a dynamics different than that of the criterion. Using Zomet et al. (2016), we explored the effects of externally added noise, and found a reduced influence of the flankers, even when sensitivity remained fixed with increasing noise by increasing the target contrast.

To explain the effects of RT, we extended the time-independent SDT explanation (Fig. 2) by using the drift diffusion model (DDM) (Gold & Shadlen, 2007; Shadlen & Gold, 2004). More specifically, we modeled the influence of the flankers as a change in the rate of evidence accumulation (drift rate), which leads to the time-independent effects on criterion that are maintained even at the slowest RTs. The time-dependent effects on criterion were modeled as a change in the starting point (Dekel & Sagi, 2020b). When contrast is increased and external noise is introduced (Zomet et al., 2016), the influence of the starting point becomes negligible, whereas the influence of the drift rate remains about the same,

consistent with an increase in gain making an additive offset negligible. In Zomet et al. (2008) the observers are slow (because of age or lack of practice); thus, the criterion effects are much smaller, showing little dependence on RT, if at all. This agrees with our previous results showing that individual differences in TAE can be explained by RT differences (Dekel & Sagi, 2020a). Although additional modeling details remain somewhat speculative, the basic idea used here, of separately considering RT-dependent and RT-independent effects and interpreting them as changes in how evidence is interpreted (the drift rate) and what priors exist before the evidence (the starting point) seems quite robust.

The presented RT analysis provides further insight into processes underlying statistical decisions. Of particular interest are limitations imposed on such decisions. We identified here a limit on the number of statistical distributions that can be efficiently learned by observers facing decisions on stimuli sampled from different distributions. Considering the example presented in Figure 2, limitations on the internal representations of response distributions are expected to be universally expressed, i.e., to introduce biases in slow responses. Indeed, distance-dependent biases are shown in all experimental conditions except two (see Figure 9). An experimental condition not showing slow effects is the one with the external noise amplitude approaching the threshold (noise amplitude=0.75threshold). We attribute the reduced distance effect to the increase in the False Alarm rate (otherwise present only at target locations near the flanker) at all distances caused by the external noise approaching the threshold. Of critical importance for the present theoretical framework is the reduced effect in the "Fix" condition of Polat and Sagi (2007). Here, unlike in the "Mix" condition, the different target-flanker distances are blocked, leaving us with only two internal distributions to learn (Signal and Noise); thus, the imposed constraints do not apply. Within the context of the DDM framework, this effect can be quantified as drift rate asymmetry (Model 2-4, Fig. 8, Fig. A2), which is the sum of the up (target present) and down (target absent) drift rates ($\nu_+ + \nu_-$), plotted in Figure 9. For an unbiased observer, this sum is expected to be zero. Figure 9A presents drift-rate asymmetry in the fitting results (Fig. A2) for both the "Mix" and the "Fix" conditions of Polat and Sagi (2007), showing a marked difference between conditions. In contradistinction, biases due to shifts in the diffusion starting point are very similar in the "Mix" and the "Fix" conditions (Fig. 9B). Accordingly, we attribute the differences in results between the "Mix" and the "Fix" conditions to objective factors affecting decisions, i.e., to differences in the evaluation of the internal responses, rather than to subjective factors such as priors (see the Introduction). We can conclude that the biases in the "Mix" condition can be explained by the distance-dependent excitatory effects of the flankers, operating in both target present and absent trials, underlying the perceived filling-in.

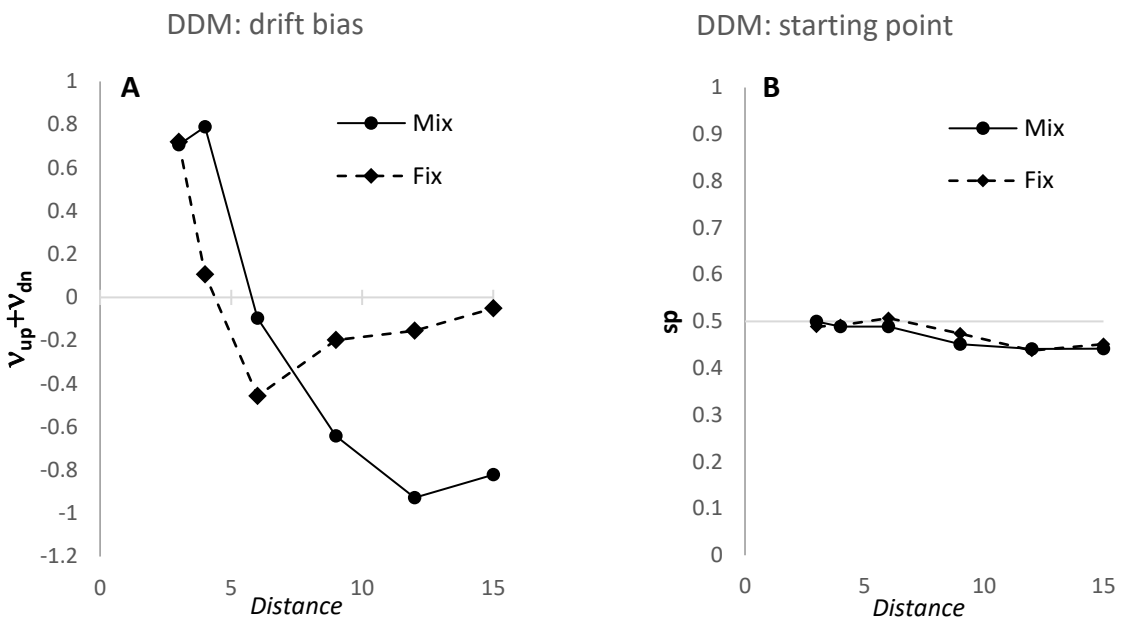

Figure 9: *(**A**) Drift rate asymmetry, quantified as the sum of the up and down drifts, evaluated by fitting Model 4 (Fig. A2) to the data of Polat and Sagi (2007). A positive bias leads to increased Hit and FA rates, and thus, to lower decision criteria. Note the difference between the two conditions (see Fig. 1). (**B**) Drift starting point (sp), evaluated by fitting Model 4 to the data of* Polat and Sagi (2007). *Biases due to prior information or payoff are indicated by deviations from sp=0.5. Both experimental conditions show unbiased sp at short distances but a lower sp at longer distances, resulting in lower Hit and FA rates at these distances (higher criteria).*

Acknowledgments

This research was supported by the Basic Research Foundation, administered by the Israel Academy of Science (grant # 2494/21, UP and DS).

References

Amiaz, R., Zomet, A., & Polat, U. (2011). Excitatory repetitive transcranial magnetic stimulation over the dorsolateral prefrontal cortex does not affect perceptual filling-in in healthy volunteers. *Vision Research*, *51*(18), 2071-2076. https://doi.org/10.1016/j.visres.2011.08.003

Chen, C. C., & Tyler, C. W. (2001). Lateral sensitivity modulation explains the flanker effect in contrast discrimination. *Proceedings of the Royal Society B-Biological Sciences*, *268*(1466), 509-516. https://doi.org/DOI 10.1098/rspb.2000.1387

Dekel, R., & Sagi, D. (2020a). A decision-time account of individual variability in context-dependent orientation estimation. *Vision Research*, *177*, 20-31. https://doi.org/10.1016/j.visres.2020.08.002

Dekel, R., & Sagi, D. (2020b). Perceptual bias is reduced with longer reaction times during visual discrimination. *Communications Biology*, *3*(1), Article 59. https://doi.org/10.1038/s42003-020-0786-7

Gold, J. I., & Shadlen, M. N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Sciences*, *5*(1), 10-16. https://doi.org/Doi 10.1016/S1364-6613(00)01567-9

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, *30*, 535-574. https://doi.org/10.1146/annurev.neuro.29.051605.113038

Gorea, A., Caetta, F., & Sagi, D. (2005). Criteria interactions across visual attributes. *Vision Research*, *45*(19), 2523-2532. https://doi.org/10.1016/j.visres.2005.03.018

Gorea, A., & Sagi, D. (2000). Failure to handle more than one internal representation in visual detection tasks. *Proc Natl Acad Sci U S A*, *97*(22), 12380-12384. https://doi.org/10.1073/pnas.97.22.12380

Gorea, A., & Sagi, D. (2001). Disentangling signal from noise in visual contrast discrimination. *Nat Neurosci*, *4*(11), 1146-1150. https://doi.org/10.1038/nn741

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. John Wiley.

Polat, U., & Sagi, D. (1993). Lateral interactions between spatial channels: suppression and facilitation revealed by lateral masking experiments. *Vision Research*, *33*(7), 993-999.

Polat, U., & Sagi, D. (1994). The architecture of perceptual spatial interactions. *Vision Research*, *34*(1), 73-78.

Polat, U., & Sagi, D. (2007). The relationship between the subjective and objective aspects of visual filling-in. *Vision Res*, *47*(18), 2473-2481. https://doi.org/10.1016/j.visres.2007.06.007

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput*, *20*(4), 873-922. https://doi.org/10.1162/neco.2008.12-06-420

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, *111*(2), 333-367. https://doi.org/10.1037/0033-295x.111.2.333

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion Decision Model: Current Issues and History. *Trends Cogn Sci*, *20*(4), 260-281. https://doi.org/10.1016/j.tics.2016.01.007

Shadlen, M. N., & Gold, J. I. (2004). The Neurophysiology of Decision Making as a Window on Cognition. *Cognitive Neurosciences Iii, Third Edition*, 1229-1241. <Go to ISI>://WOS:000279844100097

Voss, A., & Voss, J. (2007). Fast-dm: a free program for efficient diffusion model analysis. *Behav Res Methods*, *39*(4), 767-775. https://doi.org/10.3758/bf03192967

Zak, I., Katkov, M., Gorea, A., & Sagi, D. (2012). Decision criteria in dual discrimination tasks estimated using external-noise methods [Research Support, Non-U.S. Gov't]. *Atten Percept Psychophys*, *74*(5), 1042-1055. https://doi.org/10.3758/s13414-012-0269-0

Zomet, A., Amiaz, R., Grunhaus, L., & Polat, U. (2008). Major depression affects perceptual filling-in. *Biological Psychiatry*, *64*(8), 667-671. https://doi.org/10.1016/j.biopsych.2008.05.030

Zomet, A., Polat, U., & Levi, D. M. (2016). Noise and the Perceptual Filling-in effect. *Scientific Reports*, *6*, Article 24938. https://doi.org/10.1038/srep24938

Appendix

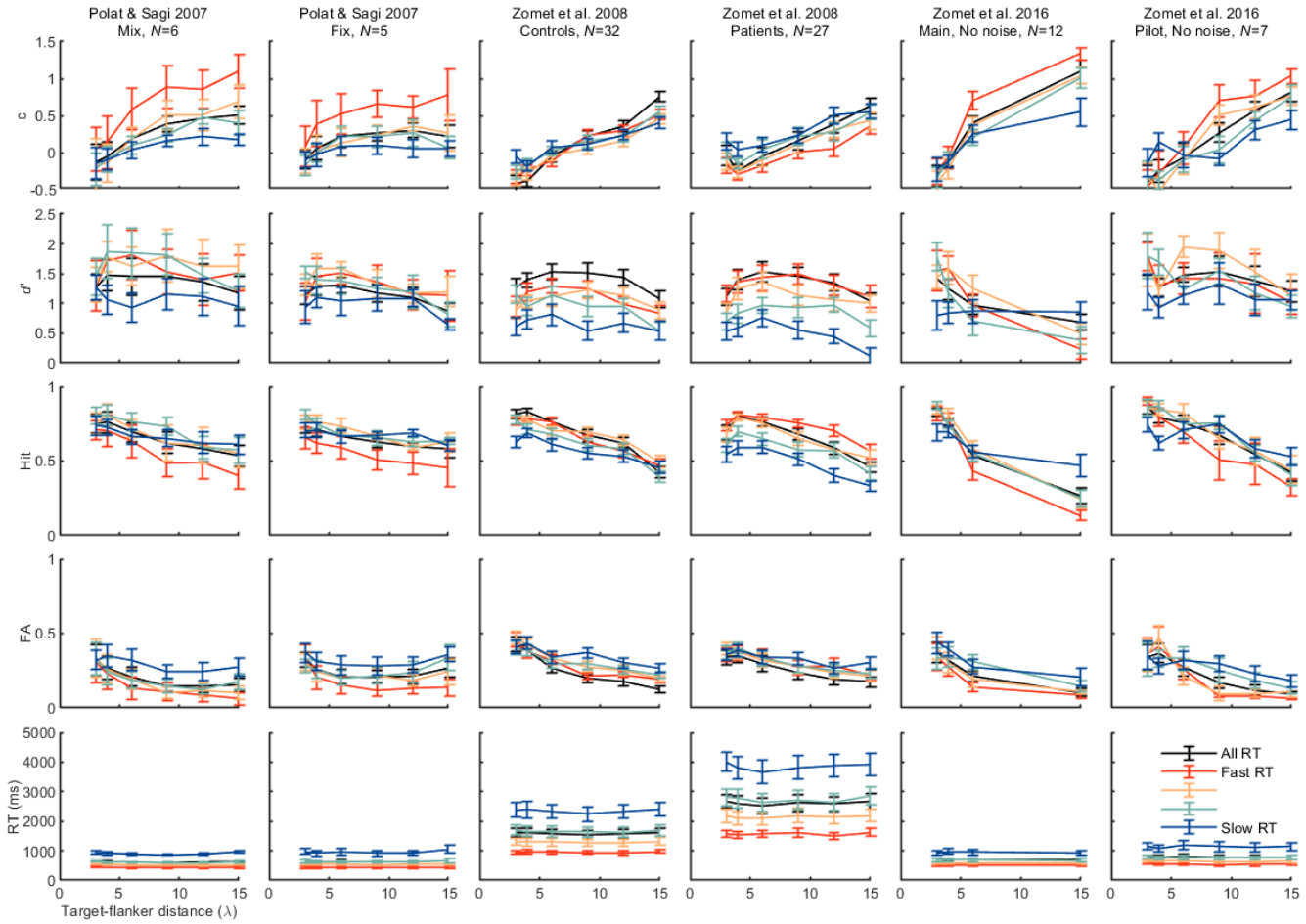

Fig. A1: Target-flanker distance and RT. *Shown, in six experiments (columns), are five behavioral measures (rows): internal criterion (c), sensitivity (d'), hit rate, false alarm rate, and RT, as a function of the target-flanker distance. Data were split into four bins (colors) by sorting the measured RTs. Error bars are ±1 SEM. Analysis of (Polat & Sagi, 2007; Zomet et al., 2008, 2016). Some data were reproduced from Figs. 1, 2, 4, and 6.*

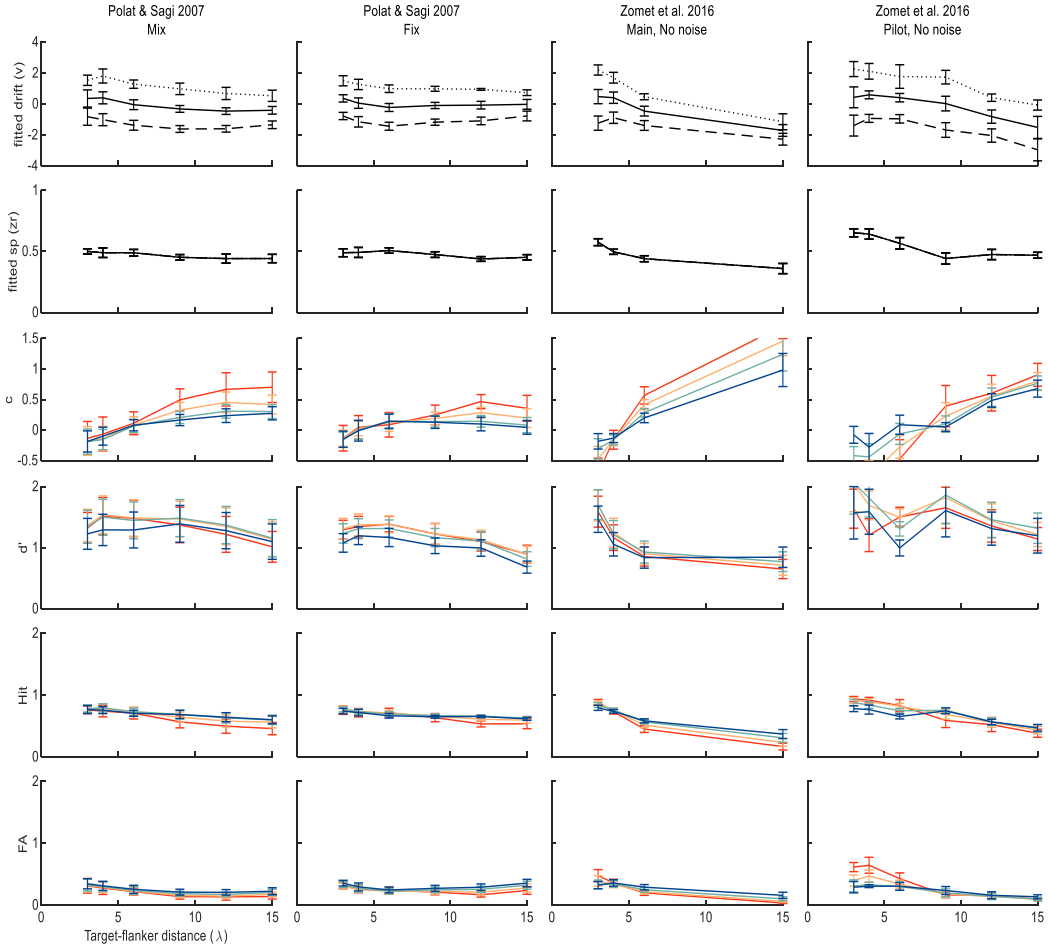

Fig. A2: Fitting behavior to Model 4. *Top two rows present the recovered parameters: drift rate (v_+, v_- and their mean) in the first row, and the starting points in the second raw, Rows 3-6 present the recovered behavior (c, d', pHIT and pFA). To obtain analytical model distributions and to fit Model 4 to behavioral data, we used the Fast-DM software with the Kolmogorov-Smirnov (KS) setting (Voss & Voss, 2007). Fitting was performed separately for each observer. The drift rate (v) was fitted separately for each combination of the target stimulus (target present/absent) and the target-flanker distance (four or six possible distances) (8 or 12 parameters); the starting point (sp) was fitted separately for each target-flanker distance (four or six possible distances) (4 or 6 parameters), and the non-decision time (t_0) and bounds-separation (a) were fitted once for all configurations (2 parameters). Overall, the model recovers the important features of the experimental results.*

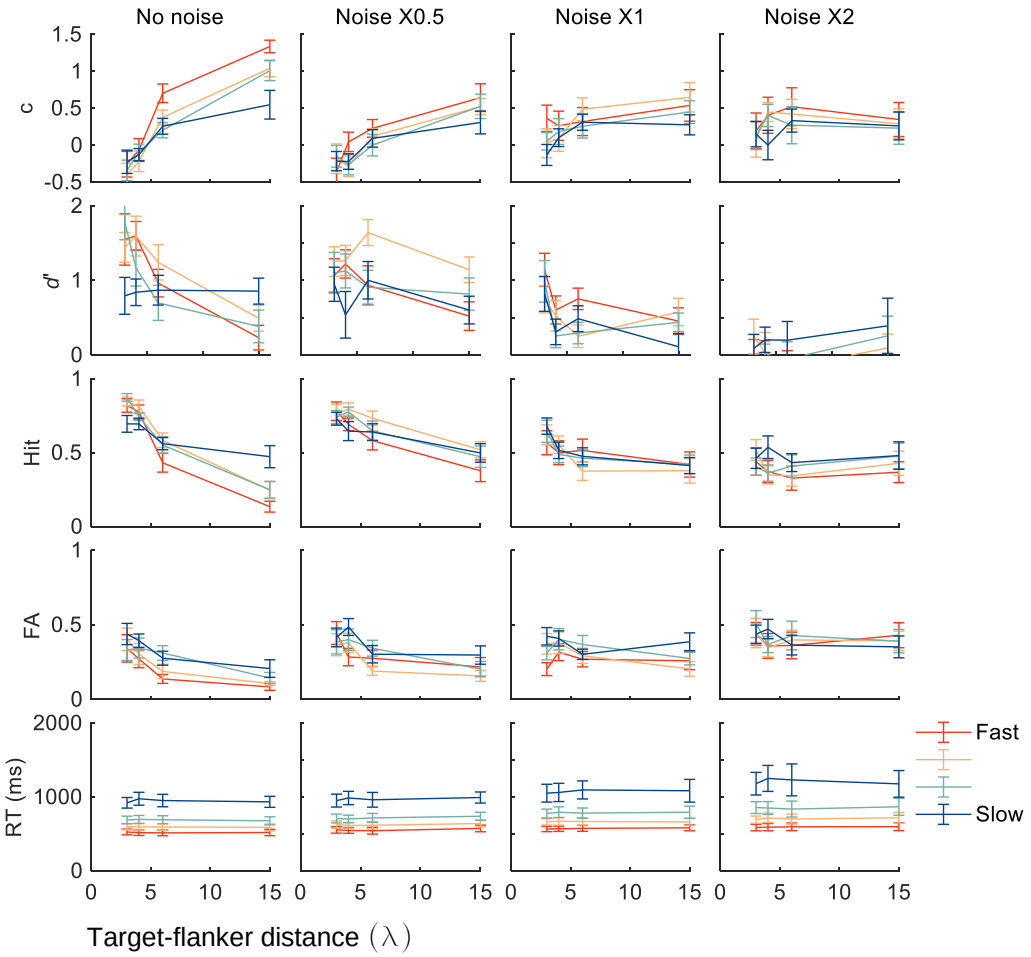

Fig. A3: RT and external noise. *RT analysis of the "main" experiment of* (Zomet et al., 2016), and the *layout following Fig. A1. The first row is reproduced from Fig. 4B.*

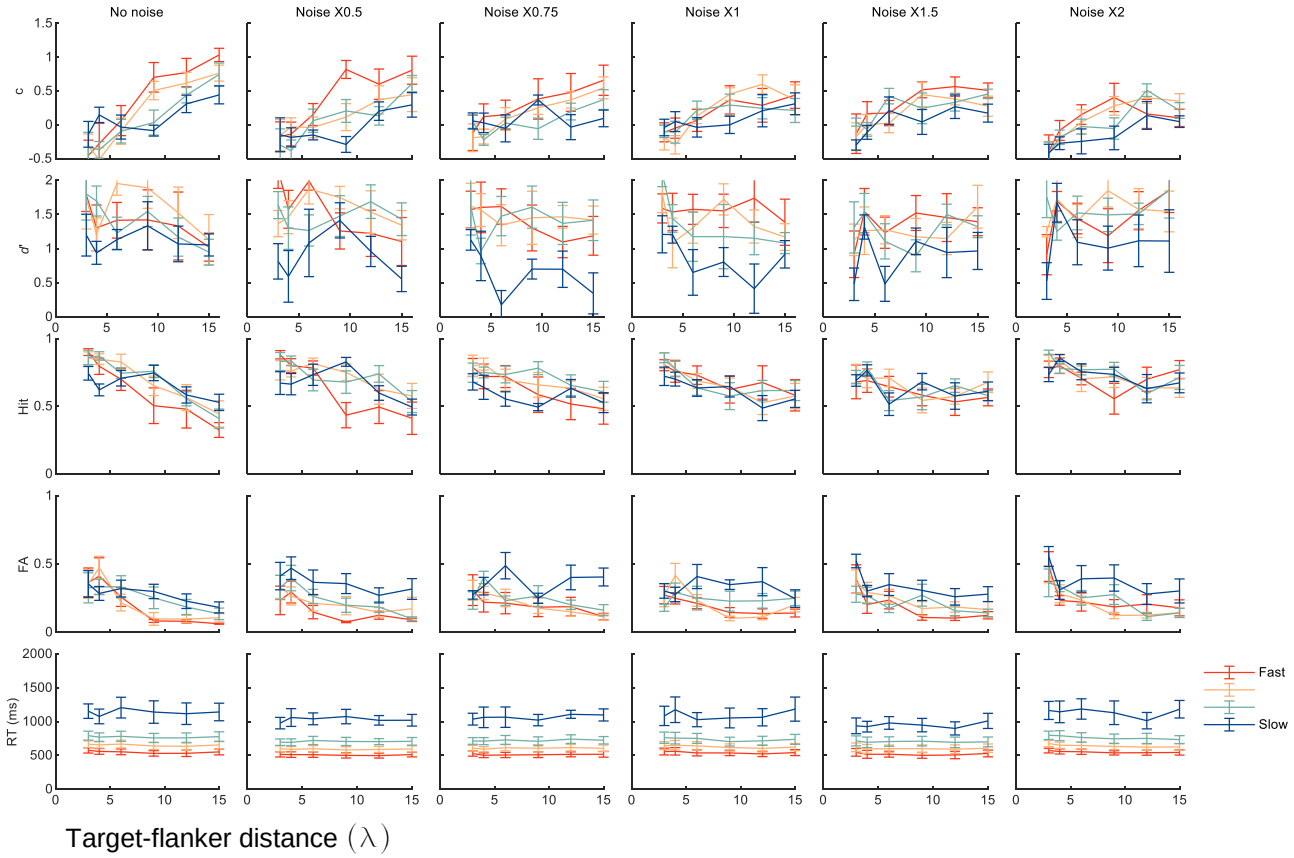

Fig. A4: RT and external noise. *RT analysis of the "pilot" experiment of* (Zomet et al., 2016). *The first row is reproduced from Fig. 4C.*

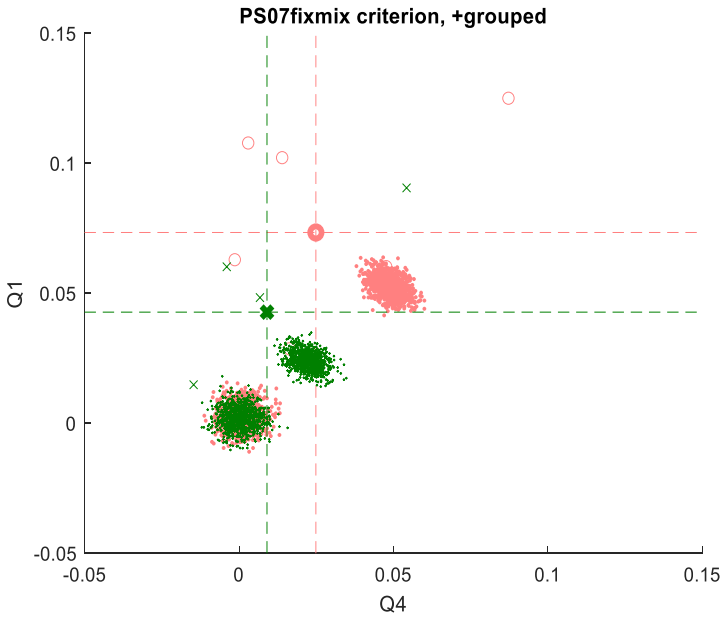

Figure A5: *Bootstrap analysis of the time dependence of the criterion-distance effect for fast (Q1) and slow (Q4) RTs, for the Fix (green Xs) and Mix (red Os) conditions measured in Polat & Sagi (2007) (see Figs. 2 and 3 in the present article). Datum points are presented for single observers (thin symbols) and their mean (thick symbols). Dot clusters represent the results of the data analysis with RT shuffling (top red and middle green clusters), and with distance shuffling (overlapping red green clusters), each with 1000 repetitions. As can be clearly seen, the RT effect (Q1 vs Q4) in both the Mix and the Fix conditions is highly significant (p<0.001). Regarding the distance effect, in examining the mean results against the shuffling cluster at the origin, there is a very high statistical significance in the Mix condition, for both fast (Q1) and slow (Q4) RTs (p<0.001), and in the Fix condition for fast RTs (Q1: p<0.001) but less so for slow RTs (Q4: p<0.01).*

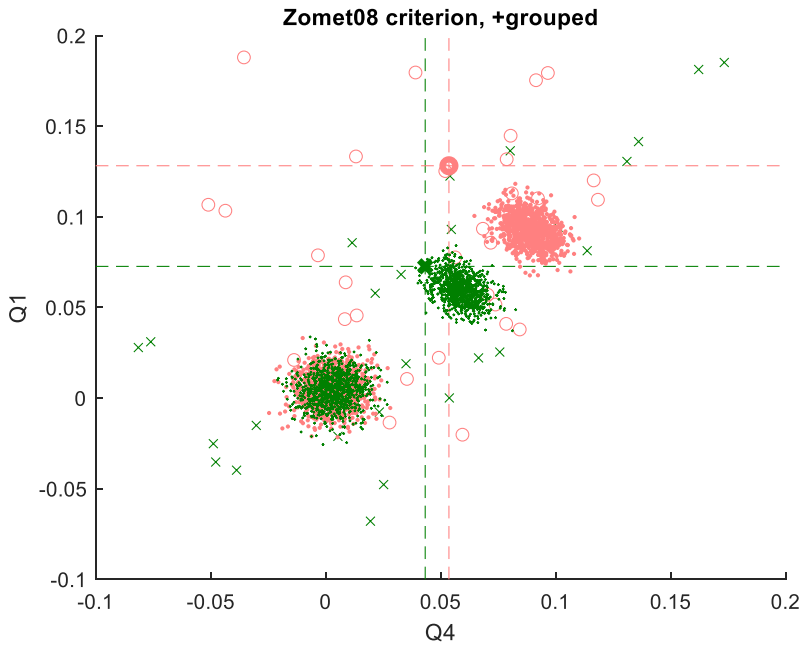

Figure A6: *Bootstrap analysis of the results from Zomet (2008). Red 'O's and green 'X's represent individual data and group means (bold symbols), for control (red) and patients (green). Clustered dots (N=1000) represent bootstrap analysis for RT shuffles (segregated red and green clusters) and distance shuffles (overlapping red and green clusters at the origin). Controls and patients show highly significant distance effects (p<0.001, relative to clusters at origin); however, only the controls show statistically significant RT effects (p<0.001). For patients, Q4 is significant at p<0.04, whereas Q1 shows p=0.28.*