Supplemental Information

Supplemental Data

Figure S1: Contrast matching experiment

To determine the apparent contrast of adapted letters, observers completed a contrast matching experiment. For 60 seconds on the first trial and 5 seconds thereafter, a stream of letters adapted one location, either 12° left or right of fixation, at random across runs. Subsequently, a *standard* letter was presented at the adapted location and a *test* letter was presented at 12° eccentricity on the other side of fixation. The standard and test letters were the same Sloan letter (randomly determined on each trial). The Michelson contrast of the standard letter was fixed at 0.8, while the contrast of the test letter varied across trials according to the method of constant stimuli.

Observers first indicated how many letters they saw (0, 1 or 2), and then which letter was higher in contrast (left or right). To ensure observer vigilance and allow assessment of biased reporting, 10% of trials were "catch" trials in which the standard (adapted) letter was not presented. Observers showed little bias to over-report the number of letters they perceived, with three out of four responding "2" on less than 6% of catch trials (where only one letter was shown). The fourth observer (naïve N2) over-reported letters on 12% of catch trials. Each observer completed at least three runs of 100 trials (10 trials at each contrast level for both gradual and abrupt onsets per block), for both adapted and unadapted standard letters.

Contrast responses on valid trials (where observers accurately reported seeing two letters) were pooled over blocks and fitted with a cumulative Weibull function using bootstrapping software, the 50% point of which was taken as the Point of Subjective Equality [1, 2].

When two letters were shown and both were reported (73.4% of non-catch trials), standard letters that were physically 0.8 contrast appeared ~0.2 contrast after adaptation (Figure S1). On 26.6% trials (Figure S1C white bar), the 0.8 contrast standard letter was not seen at all: its perceived contrast was lower than its contrast detection threshold. This high frequency of complete disappearances was not predicted from the psychometric function of apparent contrast attenuation following adaptation. Observers saw test stimuli of 0.05 contrast on more than 90% trials (Figure S1B open circles), therefore detection threshold (75%) is lower than 0.05. If complete disappearances of the standard letter simply reflect apparent contrast reduction to below contrast detection threshold, the standard letter should disappear on fewer than 15% of trials (a conservative estimate since detection threshold is below 0.05 contrast), not the 26.6% observed. This means that the mechanism of disappearances in AIB is separable from the contrast reduction caused by adaptation.

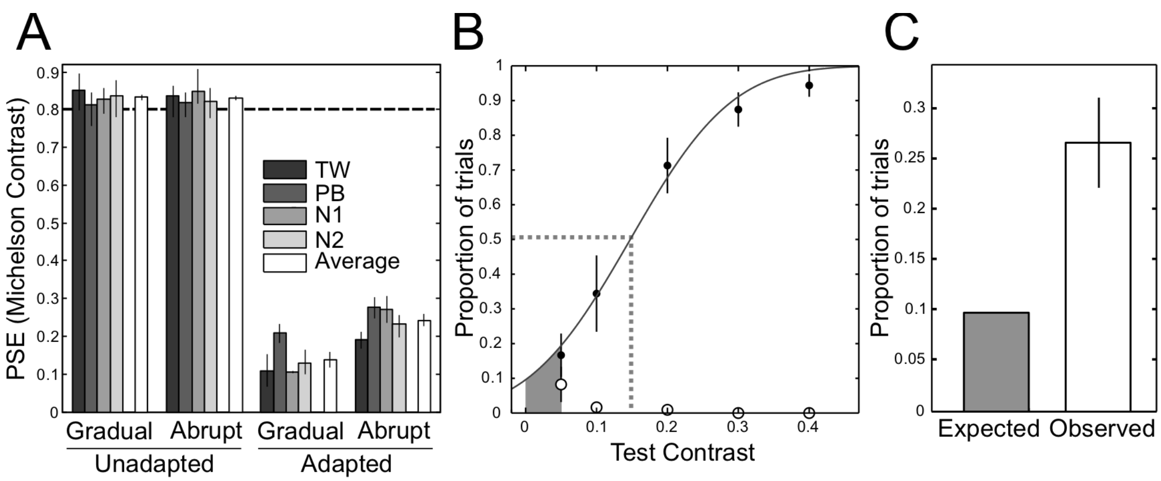

Figure S1. Contrast matching experiment

(A) The PSE between a standard letter at 0.8 Michelson contrast (dashed line) presented gradually or abruptly, with or without adaptation, and an unadapted test letter of variable contrast, on trials where the observer reported seeing two letters. Each observer (grey bars) and the average (white bars) match are shown. Error bars depict 95% confidence intervals on the PSE obtained via bootstrapping, except on the average data point, where error bars depict +/- 1 SEM between observers. (B) Adapted gradual condition, averaged across observers (error bars show +/- 1 SEM). Filled circles show the proportion of "test contrast higher" responses for trials where both letters were seen, the solid curve shows the fitted cumulative Gaussian and the dashed

line shows the PSE. Open circles show the proportion of trials where test stimuli were not seen. If disappearances were predicted simply from apparent contrast reduction to below contrast detection threshold, they would be expected to occur less than 10% of the time (grey region), assuming a conservative detection threshold of 0.05. (C) Observed disappearances occur nearly three times as frequently as expected from apparent contrast alone.

Figure S2: Logistic model fits to identification data

To quantify our observations from Figure 3, that the perceived number of letters seemed a better predictor of identification performance than the number physically present, we conducted separate logistic regressions on the identification data: one with the physical number of letters as a predictor, the other with perceived number as predictor. We quantified how well these models predicted performance by generating Receiver Operator Characteristic curves and comparing the area under the curve. This analysis was performed on all trials pooled across all observers. Note that this included trials in which observers reported seeing no letters. These "zero reports" represent very few responses: 2 trials in the unadapted 0.8 flanker abrupt condition and 5 trials in the adapted gradual condition.

In addition, the logistic regression analysis included trials where observers reported seeing more letters than were physically present in the display. While it is interesting to consider whether crowding is caused by the perception of illusory features, these responses were too infrequent to allow a robust assessment of this suggestion (over-reports occurred on 2.2% of all trials). We searched for a trend in our data by pooling all trials across all observers, then compared the performance on trials where observers reported more letters than were physically present to performance when observers reported the veridical number of letters. In 62.5% of 24 cases, performance was worse when observers over-report the number of letters perceived. While this result could suggest that illusory features may cause crowding, a position that fits with our proposal that crowding depends on the perceived number of features, we do not have enough data to support this claim statistically.

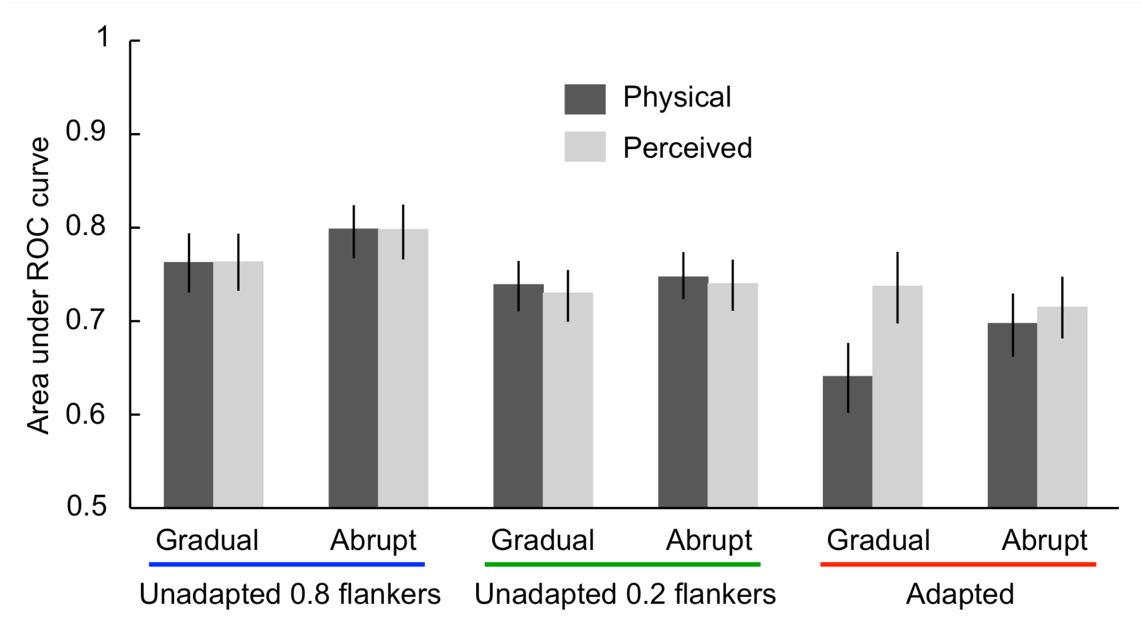

Figure S2. Logistic model fits.

Bar plot of the area under the Receiver Operating Characteristic (ROC) curve for logistic regression models predicting correct identification with perceived and physical number of letters as predictors, for all conditions. Error bars show 95% confidence intervals obtained via bootstrapping. In the adapted gradual condition, the number of letters perceived is a better predictor of correct letter identification than the number of letters physically present. This is not the case for any other condition. Excluding over-reports or zero-reports made no difference to the outcome of this analysis.

As can be seen in Figure S2, in the adapted gradual condition, letter identification performance was better predicted by how many letters observers perceived on a given trial, rather than how many were physically present.

Supplemental References

1. Wichmann, F., and Hill, N. (2000). The psychometric function: I. Fitting, sampling, and goodness of fit. Perception and Psychophysics *63*, 1293-1313.

2. Wichmann, F., and Hill, N. (2000). The psychometric function: II. Bootstrap-based confidence intervals and sampling. Perception and Psychophysics *63*, 1314-1329.