Gestalt Grouping Leads to Reduced Flank Interference in Vernier Tasks

Keenan May
Rice University
krm2@rice.edu

ABSTRACT
Vernier Acuity has been shown to be negatively effected by crowding flanks. The present study aims to address the issue of whether a flank configuration that obeys certain Gestalt principles can either provide an increase in acuity or have a quantifiable reduction in interfenerence when compared to similar, but rule-breaking flank configurations.

INTRODUCTION
Vernier acuity is a classic example of what is known as hyperacuity, which refers to an ability for finer discrimination then the human perceptual system, viewed as a purely bottom-up system, should be capable of (Westheimer and Hauske, 1975). Vernier acuity is well-known phenomenon, and was first documented by Hering (1899). The maximum resolution achievable by cones in the fovea is around 1 arcminute in terms of positional discrimination. However, for certain kinds of tasks, the human brain is capable of discriminating spatial offsets as small as 1-5 arcseconds (Klein and Levi, 1985). Vernier acuity is a classic example of such a hyperacuity (see Westheimer (2009) for a comprehensive review .) It refers to the ability to discriminate the offset direction of two vertical lines, or, more commonly, to determine whether such lines are collinear (depending on the experimental paradigm used). Of note to the present study is the fact that performance on this task is degraded severely when flanks of a certain distance are added (Westheimer and Hauske, 1975). With no flanks, vernier acuity is often found to be around one or two arcseconds, which is about $1/40_th$ of a cortical perceptive 'hypercolumn' . Hypercolumns are proposed cortical devices which facilitate 'magnification' of stimuli close to and inside the fovea, and provides a framework for understanding hyperacuity and why flanks of a certain distance might effect it (Levi et al., 1985). It was originally documented by Westheimer & Hauske (1975) that vertical flanking lines introduced either concurrently or with a stimulus onset asynchrony of less than 100 ms at lateral distances of 3-10 degrees from the vernier lines can almost triple the vernier threshold . For the subjects tested, the distance of greatest flanker effect was around 2-4 degrees from the target. This was found to vary by subject to some extent; as such, the standard experimental method for studying flanker effects involves calculating this optimal distance on a per-subject basis (Sayim et al., 2010) .

THE NATURE OF THE FLANKING EFFECT
The fact that distractor flanks can interfere with such a fine discrimination at relatively large distances is a matter of some mystery, and has seen several competing explanations (Levi et al., 1985). Badcock & Westheimer (1985), building on the work of Westheimer and McKee (1977), suggested an explanation based off of spatial pooling. This explanation is based off of the principles of information aggregation and the way flanks exhibit perceptual repulsion and attraction on target stimuli. Badcock & Westheimer (1985) observed that flanks within the optimal distance (that is, flanks that fall in the inhibitory edges of hypercolumn receptive fields) that have positive contrast polarity tend to 'pull' the target toward the flank; conversely, when flank contrast polarity is negative, the target is 'pushed' away from the flanks. This was theorized to be due to the fact that the perceptual system assigns the position of an object to the centroid of a luminance distribution that represents the raw input from about 3 degrees around the object being foveated. In theory, this would cause the presence of a close flank with equal contrast polarity to 'pool' with the input from the target line, thus shifting the centroid of the luminance distribution toward the flank. Conversely, when the contrast polarity is reversed, the pooling paradigm would suggest that the centroid of the luminance curve would be shifted inward -away from the flank, thus increasing difficulty and worsening acuity-, which was exactly what Badcock and Westheimer observed.

CONFIGURATION AND CONSEQUENCE
Another potential explanation for the flank effects centers around the notion that perceptual grouping is important in determining the position of objects in the fovea. In this paradigm, it is more than just the raw *position* of luminance nodes; instead the perceptual system is performing configuration-based logic in order to determine object position. Malania et al. (2007) tested the influence of various flank configurations on vernier acuity, and found that contextual grouping had a large impact on flank effectiveness. The investigators drew different conclusions depending on the kind of configuration used. Overall, of note was the observation that flank interference was in no way a linear summation of the interfering capabilities of each flank *component*. Acuity without flanks was observed to be about 11 arcseconds (due to characteristics of the particular experimental method used). Their observations, which have particular import with regards to the present study, were as follows. First, flanking lines equal in height to one of the vernier lines were found to have optimal interference when there were about two on each side, with performance degradation sloping off to around a 2x vernier threshold increase. Second, flanks closer to the same length as the summed height of the vernier lines yielded more interference than longer or shorter flanks. Finally, irregular flank configurations caused more interference. Of note is

that even the most benign flanking configuration did not *facilitate* vernier discrimination. There was always at least a 40% increase in threshold offset over the unflanked threshold. However, the threshold for verniers presented inside a series of tall vertical flanks was actually significantly *lower* than the threshold for verniers presented between only two such flanks, indicating that Gestalt grouping may have a role in positional judgment. These findings were followed up on by Sayim et al. (2008), in which the investigators isolated several additional grouping-based flank properties. Of particular note was the observation that a condition in which a series of vertical flank lines that got taller with increased lateral distance from the vernier targets produced the same threshold as did a set of lines with the same quantity, location, and contrast polarity polarity, but that remained constant in height. In addition, conditions in where there was a color contrast between flanks and verniers leads to improved acuity over a lack of contrast, indicating that pure luminance cannot fully explain the flanking interference effect (the investigators used control studies to verify this). Finally, the investigators found evidence that popping the vernier out from the flanks using stereoscopic depth decreased offset thresholds, indicating that relevant perceptual grouping occurs at a higher level than binocular fusion. The conclusion drawn by Sayim et. al. (2008) was that at least part of the nature of vernier interference occurs at a high level in the visual system. Saarela et al. (2010) used a Gabor task to study the effectiveness of the spatial pooling theory in explaining observed task interference. They found that grouping the flankers (rotated 'T' characters, in this case) allowed conditions in which there were twice as many flankers occupying the same physical space to cause significantly less interference. Their conclusion was that when flankers experienced regular open spacing, a homogenous texture was formed with the target stimulus, thus making it difficult to make judgments about the target.

One largely unstated conclusion is the potential to use Gestalt principles to generate flanks that actually facilitate target judgments in the vernier task. The potential for flanks to actually have a threshold-lowering effect has not been explored at length in the literature. The ability of 'noise elements' that have perceptual Gestalt configurations to *improve* detection speed and accuracy of desired targets has been applied to a variety of tasks, such as parallel character search(W.P. Banks, 1976). Sayim et. al. (2010) does not investigate this potential directly, but it does lay out a framework for using the *degree* to which Gestalt grouping reduces vernier stimulus deterioration can be used to quantify and investigate the relative strengths of Gestalt laws. The investigators used a series of flank configurations, some of which obeyed Gestalt laws and some of which did not, and examined the resulting interference with vernier offset direction discrimination. They found that flanks forming 3D cuboids exhibited almost the same level of interference as did simple rectangles, even though the former flank was generated by *adding* five lines to the latter flank. Both of these conditions led to better performance compared to the basic single vertical line flank, which also supports the relevance of Gestalt factors, since, again, the rectangles and 3D cuboids were generated by adding lines (and presumably complexity and task noise) to the basic flank condition. Finally, Sayim et. al. (2010) found that taking the same lines found in the 3D cuboid condition and scrambling them (while keeping proximity generally the same) led to vastly decreased acuity. All of this evidence supports the supposition that grouping has an influence on acuity interference. The authors subsequently propose that vernier offset threshold levels can be used as a metric for investigating Gestalt factors. The present study is aimed both at replicating the results of Sayim (2010) and beginning to perform such quantitative analyses on Gestalt principles.

GOALS

The present study was aimed at addressing several questions raised in response to Sayim et. al. (2010). The first is the methodological question of how to best get accurate threshold readings given limited equipment, software, and subject time. The aim of these endeavors was to open up the study of Gestalt factors in this fashion to more researchers, and to provide a methodological template for how to do so. The second question was whether or not Gestalt grouping of flanks can be used to produce a facilitative effect with regards to vernier acuity, as opposed to the traditional detrimental effect. The third was to provide information for the development of an objective metric of Gestalt feature 'strength' based on the level of acuity interference over related, but non-grouping, conditions.

STIMULI TYPES

A variety of stimuli were selected for use, each aimed at evaluating the influence of a different Gestalt principle. The first stimuli item was the simple Vernier task involved two near-collinear vertical lines, with no flanks (Fig. 1). This was included as a control condition to provide a meaningful comparison metric. The second stimulus condition (Fig. 2) was a set of simple line flanks of the same length as the two verniers and the gap between them. These flanks were placed statically at a distance of .1 minutes of arc. It should be noted that the original intent was to determine the optimal flank distance using the procedure described by Sayim (2010) , in which the proximity of the flanks was adjusted until the acuity threshold was twice that observed under the baseline unflanked condition. However, this procedure proved unwieldy and time-consuming, and eventually we chose to simply set the flank distance to one of the common distances used in Sayim (2010). The next two stimulus (Figs. 3-4) types

Figure 1. Unflanked Vernier Condition

Figure 2. Flanked Vernier Condition

were aimed evaluating the influence of the basic Gestalt principle of closure on Vernier acuity. These can be described as outward-pointing isosceles triangles, either with edges that meet or edges that do not. The expectation was that acuity would be better for the condition in which the three lines formed a contiguous triangle versus the condition in which the same three lines did not complete the form. The next

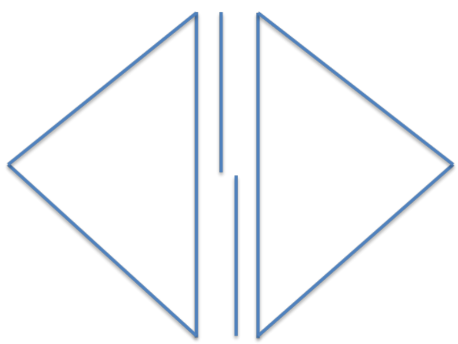

Figure 3. Flanked Vernier Plus Triangles Condition

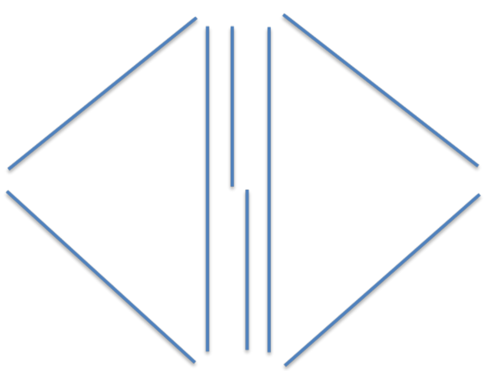

Figure 4. Flanked Vernier Plus Broken Triangles Condition

three conditions (Figs. 5-7) were designed to address two questions. The first is whether vernier acuity remains at hyperacuity levels given extreme vertical target distance. The second is whether the presence of illusory contours (Victor and Conte, 2000) facilitates or inhibits such performance. In order to evaluate the question about vertical vernier distance, we first tested a condition involving no flanks, but a widely separated pair of verniers. The other two stimuli in this group constituted a group of three lines on each side, equal in length to the verniers. In the illusory contour condition, these lines were arranged in a format that caused the interior terminator to induce the perception of a false white circle.

Figure 5. Unflanked Vertical Gap Condition

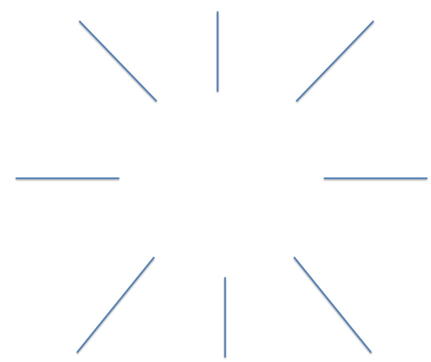

Figure 6. Illusory Contour Condition

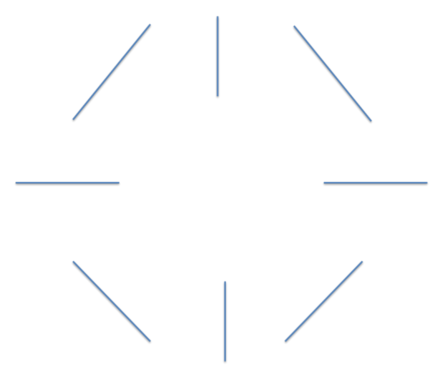

Figure 7. Broken Illusory Contour Condition

The final set of stimuli (Figs. 8-11) were designed to investigate the strength of several basic emergent features on vernier acuity. These conditions were built by displaying a single horizontal line (half the length of the vernier targets) at various horizontal distance from the bottom edge of top target. Performance was expected to differ between the 'near line' condition and the two abutting line conditions due to the presence of emergent features in the latter conditions.

Figure 8. Emergent Feature Far Line Condition

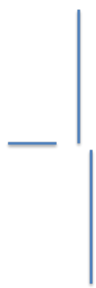

Figure 9. Emergent Feature Close Line Condition

Figure 10. Emergent Feature L Junction COndition

Figure 11. Emergent Feature T Junction Condition

DEVELOPING A METHODOLOGY

Psychopy

Since this study required an adaptive experimental procedure, and since one of the goals was to endeavor to perform the study at minimum cost, we used the open source Python module Psychopy, which was created to make performing psychophysics on everyday hardware feasible (Pierce, 2007), to present the stimuli and produce experimental output. This software package generally worked very well, and it is suitable for use in psychometric threshold experiments- particularly those in which precise display time is not a crucial factor. As such, we recommend the Psychopy package for psychophysicists endeavoring to perform studies similar to the present one.

Adaptive Staircases and Threshold Calculation

We tested multiple methods for determining the acuity threshold, which was defined in Sayim et al (2010) as the horizontal gap size at which subjects could successfully identify the direction of the offset 75% of the time . The first was a simple walking offset that followed the following logic: Repeat an offset size for *n* repetitions. If the successful detection rate is below 75%, make the offset smaller by a fraction of the previous increment. If it is above 75%, make the offset larger by a fraction of the previous increment. This method proved cumbersome and time-consuming in reaching the desired 75% detection threshold, since a large *n* was required to achieve reliable results. In order to alleviate these issues, we moved on to a second model, which made a fixed number of vernier gap size 'samples' and then generated a Probit model (since the task is a forced binary choice of left or right) capable of directly predicting the threshold necessary to achieve the desired hit rate. This method proved too inconsistent for use for this particular task, because it was difficult to get large quantities of data in the informative range between 100% hit rate and 50%. It is possible that this method would work with an improved stepping algorithm. Finally, we investigated the use of a minimum-reversals staircase employing decibel scale stepping as implemented in PsychoPy by (Pierce, 2007). This staircase was defined by a X down, Y up control logic that would raise and lower the offset gap size when X correct answers or Y incorrect answers were recorded at a certain gap size level. A reversal is defined as a point at which the direction of the up/down test changes from previous instances; theoretically, if this occurs, a point has been reached at which the task is either slightly too hard or slightly too easy for performance to be at threshold levels. After a number of these reversals have occurred, it is likely that the subject is close to the threshold implied by the values of X and Y. It should be noted that these values can only home in on certain percentage thresholds; 71% and 79% are possible, but not 75%. For the present study we employed a '1 up, 2 down' setting, which should have caused subjects to approach 71% accuracy. This method worked well with the following parameters starting off set of 2.4 arcminutes, decibel scale stepping, and a minimum of 75 trials and 10 reversals.

Display of Stimuli

Finally, the display of stimuli presented a challenge. Without an analog oscilloscope, we ran into problems with the pixel resolution of the display not being dense enough to properly present slight offset differences. In order to rectify this, we used a large display at high resolution and high distance. This allowed us to present stimuli properly. It should however be noted that the significant distance requires subjects who wear corrective lenses for nearsightedness to use said lenses during the experiment, which can lead to reported eye fatigue. Thus, it is important to give subjects ample opportunity to take breaks.

METHODS

Participants

A total of nine Rice University undergraduates participated in the study.

Apparatus

We used a high definition Sony television to display stimuli. The device was 40 cm across, and set to display at a progressive scan resolution of 1920x1080 pixels. The display was driven by a laptop running the PsychoPy software, outputting through an HDMI cable. The contrast ratio was adjusted to be high, but was not recorded. Subjects sat at a distance of 15 feet from this display, and used a Logitech wireless keyboard to provide response input.

Stimuli

Stimuli were presented as earlier described. After several conditions, subjects were prompted to take a break and then return to the test when ready.

Procedure

The procedure laid out in Sayim et. al. (2010) was followed when possible. Vernier stimuli were presented as black vertical lines, each 10 arcminutes long and separated by a 1 arcminute vertical gap. The basic flanks were 21 arcminutes long, with the same line thickness (not specified by Sayim et. al., but estimated to be similar) as the vernier targets. In the present study, for the conditions in which direct line flanks of this sort were used, their distance was set at 250 arcseconds, which was the most common distance used by Sayim et. al (2010). Stimuli were presented 400 ms after a subject's previous response input, and stayed on screen for 250 ms. Subjects had unlimited time to respond by pressing either the left or right arrow key depending on their judgment of which direction the top vernier line was offset. Subjects were instructed that there would always be a direction of offset, and to do their best even if they thought they couldn't tell in which direction the line was offset. The entire procedure took between 20 and 25 minutes.

RESULTS

Conditions were divided into three sets for the analysis of results. The first constituted the normal unflanked, flanked, grouped triangles and ungrouped triangles conditions. The second group was comprised of the unflanked large vertical gap condition, the illusory contour condition, and the broken illusory contour condition. These were grouped due to the fact that the large vertical distance of the verniers made comparison between these and the normally spaced conditions impossible without some sort of normalizing procedure. Finally, the four emergent feature conditions were grouped. A general linear model was constructed of each group. Group 1 performance did not vary significantly by condition ($F(3,8) = .23$, $p = .87$). It should be noted that the lack of statistical significance under this and other conditions is most likely a consequence of not having enough subjects. The potential for significant effects given more data is indicated by the fact that the F value is less than 1. Overall between-subjects variance was significant for group one, and similarly significant for other groups ($F(1,8) = 41.22$, $p = .00$), ($F(1,8) = 56.32$, $p = .00$), and ($F(1,8) = 38.02$, $p = .00$). In a scenario with a small subject pool and a large between-subject variance, the between-subject variance tends to be underestimated in comparison to the within-subject variance (for which there is more data and less variance). Thus, we will report findings with the theory in mind that a larger sample size might lead to statistically significant trends. The need for such a large sample size was somewhat surprising, as many similar studies have used less than five subjects. The present study lends itself to the conclusion that when equipment and (accordingly) measurement accuracy and precision are constrained, a much larger sample size is needed to counteract the increased noise. In the interest of informing future research, results are reported here regardless of statistical significance given the current (limited) data set. Figure 12 represents the results for the group 1 stimuli. As expected, acuity was damaged by vertical line flanks, but the change in acuity was negligible when lines were added to make said flanks into grouped triangles. When said triangles were broken, acuity became approximately one degree worse. Again, none of these results are statistically significant, but nonetheless it is a plausible hypothesis that, given additional subjects, similar trends might become significant.

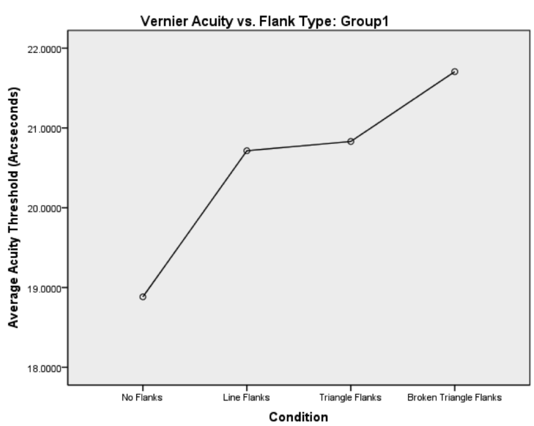

Figure 12. Average Observed Acuity Thresholds by Condition (Group 1 Flanks)

For group 2 (Fig. 13, the high distance verniers and illusory contour stimuli) , there was also no significant effect of flank type on estimated threshold ($F(2,8) = .003$, $p = .87$). While thresholds were measured to be higher for the illusory contour condition versus the broken contour condition, indicating that perhaps the contour causes the verniers to group with the flanks much in the same way that Saarela et al. (2010) described with regards to evenly spaced repeated patterns, such effects were extremely small and extremely unlikely to be significant.

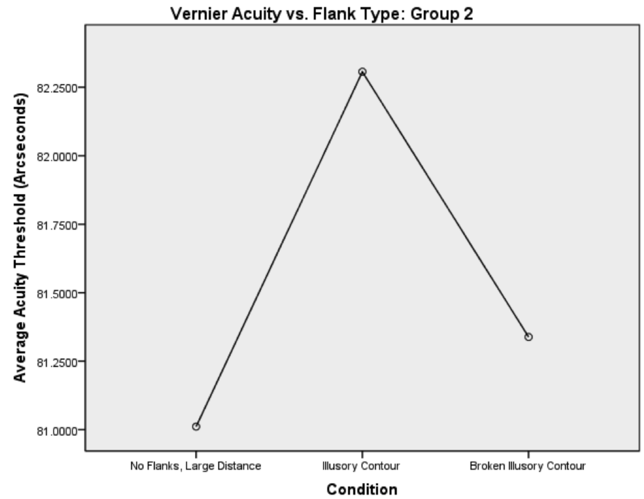

Figure 13. Average Observed Acuity Thresholds by Condition (Group 2 Flanks)

The stimuli of group 3 (Fig. 14) also elicited no significant difference by condition ($F(3,8) = 2.56$, $p = .08$)). However, (as expected), performance on the far flank condition was similar to that of the unflanked vernier task, and performance declined when the flank was moved close. There was almost no difference between the close flank and the L junction flank conditions, indicating that the presence of the emergent L junction did not significantly affect performance. Performance on the T Junction condition was about 8 degrees worse on average than the L junction condition. This particular configuration may be particularly damaging to vernier acuity because of the way it disrupts the bottom terminator by grouping with the top target line.

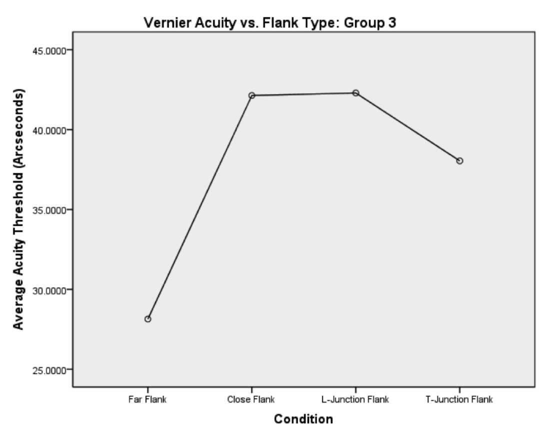

Figure 14. Average Observed Acuity Thresholds by Condition (Group 3 Flanks)

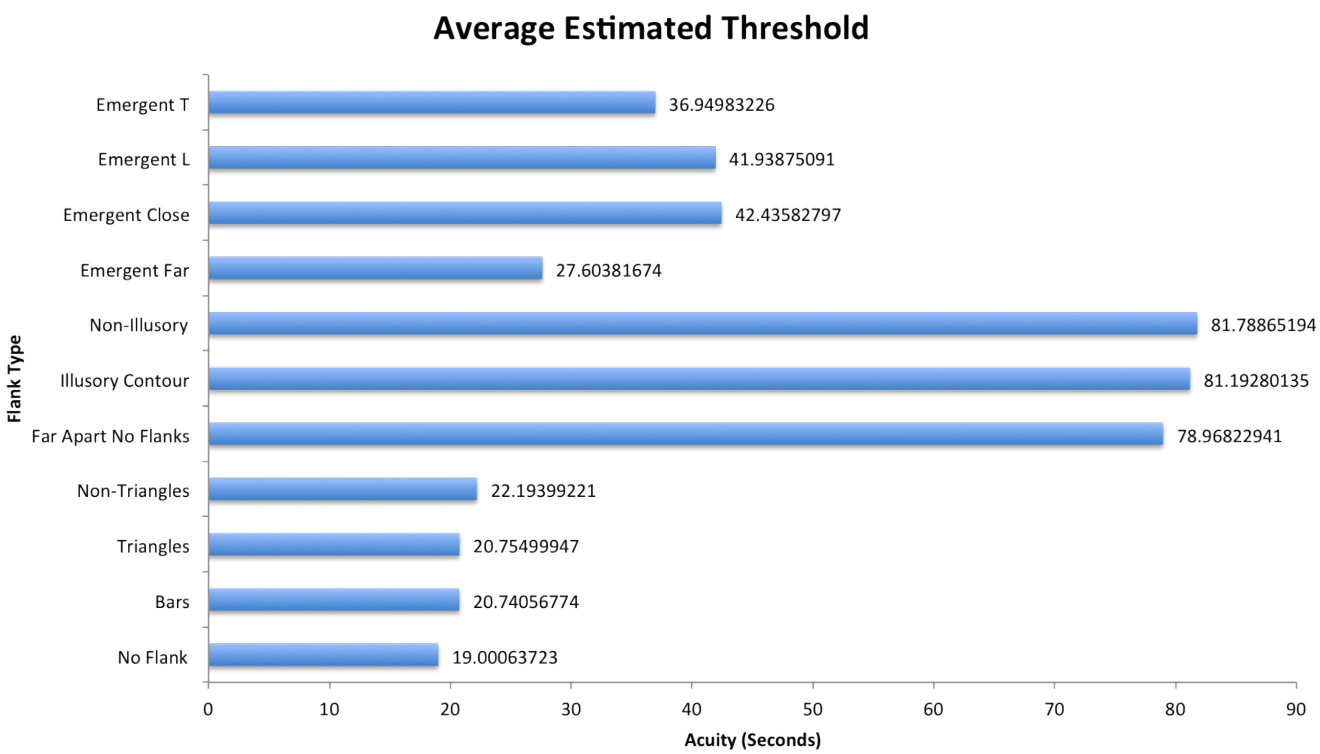

Figure 15. Average Observed Acuity Thresholds by Condition

DISCUSSION

The present study has provided a methodology for performing psychometric investigations into Gestalt strength. However, the current data set is not strong enough to allow for serious development of a numerical metric of flank strength. That said, the data does make a few basic suggestions that present further lines of investigation. First, it indicates that the effect of object closure on facilitating good perceptual grouping of the vernier targets as separate from the flanks is relatively strong. If a facilitative flank is ever found, it is likely that it will employ color contrast and simple closed flanks in conjunction with other methods of providing extremely strong perceptual grouping. In addition, there is weak evidence for the presence of illusory contour as a pattern-forming distractor. Performance was degraded when verniers were part of an illusory contour than when the same lines were angled so as not to elicit this effect. However, future studies would have to rule out the explanation that this was simply due to the number of terminators in the vicinity of the vernier stimuli. Finally, the fact that the T-junction condition caused more interference than the L junction condition leads credence to the theory that a variety of emergent features could perhaps be quantifiably described by the degree to which they interfere with vernier acuity, and that such differences may exist beyond the one noted here. The reason for this particular difference could bear further investigation, but it might have something to do with the presence of two vertical interior angles than begin to apply a more complex form for the top vernier target. In any case, further insight into the specific nature of these emergent feature effects, and the other configural effects observed here could perhaps be achieved using the methods outlined in this paper with a larger subject pool, at minimal cost and difficulty. Such research could be an important step on the road toward finally quantifying traditional Gestalt concepts.

References

Badcock, D. and Westheimer, G. (1985). Spatial location and hyperacuity: The centre/surround localization contribution function has two substrates. *Vision Research*, 25(9):1259–1267.

Hering, E. (1899). Uber die grenzen der sehscharfe. *Ber. math.-phys*, pages 16–24.

Klein, S. and Levi, D. M. (1985). Hyperacuity thresholds of 1 second: Quantitative predictions and empirical validation. *Journal of the Optical Society of America*, 2:1170–1190.

Levi, D., Klein, S., and Aitsebaomo, A. (1985). Vernier acuity, crowding, and cortical magnification. *Vision Research*, 25(7):963–977.

Pierce, J. (2007). Psychopy—psychophysics software in python. *Journal of Neuroscientific Methods*, 162(1-2):8–13.

Saarela, T., Sayim, B., and Westheimer, G. (2010). The effect of spacing regularity on visual crowding. *Journal of Vision*, 10:1–7.

Sayim, B., Westheimer, G., and Herzog, M. (2010). Gestalt factors modulate basic spatial vision. *Psychological Science*, 20:1–4.

Sayim, B., Westheimer, G., and Herzog, M. H. (2008). Contrast polarity, chromaticity, and stereoscopic depth modulate contextual interactions in vernier acuity. *Journal of Vision*, 8:1–9.

Victor, J. and Conte, M. (2000). Illusory contour strength does not depend on the dynamics or relative phase of the inducer. *Vision Research*, 40:3475–3483.

Westheimer, G. (2009). Encyclopedia of neuroscience. In *Encyclopedia of Neuroscience*, volume 5, pages 45–50. Elsevier.

Westheimer, G. and Hauske, G. (1975). Temporal and spatial interference with vernier acuity. *Vision Research*, 15:1137–1141.

W.P. Banks, W. P. (1976). Configural effects in visual information processing. *Perception & Psychophysics*, 19:361–367.