Multi-inducer grouping for curve completion: Perceptual and computational exploration

Nir-Cohen Gal
Department of Brain and Cognitive Sciences, Ben Gurion University of the Negev, Be'er-Sheva, Israel
Zlotowski Center for Neuroscience, Ben Gurion University of the Negev, Be'er-Sheva, Israel

Arad Boaz
Department of Computer Science, Ben Gurion University of the Negev, Be'er-Sheva, Israel
Zlotowski Center for Neuroscience, Ben Gurion University of the Negev, Be'er-Sheva, Israel

Ben-Shahar Ohad
Department of Computer Science, Ben Gurion University of the Negev, Be'er-Sheva, Israel
Zlotowski Center for Neuroscience, Ben Gurion University of the Negev, Be'er-Sheva, Israel

The human visual system excels in object recognition and scene interpretation even in scenes in which some (or even all) observed objects are partially occluded or fragmented. This highly efficient capacity is facilitated by constructive processes of contour completion between inducers to yield the perception of whole objects across gaps. A fundamental problem of the process is when and how the visual system groups different inducers in the visual scene between which completion occurs. Previous studies on this grouping problem, inspired mostly by relatability theory (Kellman & Shipley, 1991), focused on one good continuation condition that dictates whether a given pair of inducers would group together or not. Left open, however, was the question of how good continuation interacts with other grouping factors and how the perceptual system groups inducers when more than two of them are present in a scene (as would be in most typical natural stimuli). Here, we address both issues by exploring observers' perceptual response in multi-inducer scenes in which inducers and their relationships are determined by three independent factors: relative proximity, relative orientation, and curvature. Employing the dot localization paradigm, we indirectly inferred the type of inducer grouping constructed by the observer's visual system and analyzed the three-dimensional response function. Furthermore, we propose a simple parametric model that suggests an opponency relationship between two of the grouping factors and predicts subject's likely response with high accuracy. We also analyzed the discrepancy between our results and model the predictions made by the classical relatability theory.

Introduction

Visual object recognition is a fundamental perceptual capacity that is facilitated by rapid integration (Potter, 1976) of representations, memory, and measurable visual features (e.g., color, texture, contours) that project on the observer's retina (Biederman, 1987; DiCarlo, Zoccolan, & Rust, 2012). Because natural physical scenes typically contain several objects that require recognition before further decision making and physical interaction, the possibility that certain objects partially occlude others is the likely event rather than the accidental one. To recognize such partially occluded objects from their visible retinal fragments, our visual system performs perceptual completion (Beck & Palmer, 2002; Schumann, 1904, in Coren, 1972; Reisberg, 2013; Wagemans et al., 2012) resulting in the perception of whole and coherent objects.

Often, the visual completion problem can be reduced to a one-dimensional curve completion task performed on object boundaries (while ignoring other features of the object). In these cases, the completion happens by "interpolating" complete contours from pairs of *inducers*—the observable (and thus measurable) visual

Citation: Gal, N.-C., Boaz, A., & Ohad, B.-S. (2017). Multi-inducer grouping for curve completion: Perceptual and computational exploration. *Journal of Vision*, *17*(9):8, 1–15, doi:10.1167/17.9.8.

doi: 10.1167/17.9.8 Received September 17, 2016; published August 16, 2017 ISSN 1534-7362

information at the points at which the occluder meets the occluded object (Prazdny, 1985). In a typical scene, the number of inducers can be large, implying that this general problem of contour completion should be divided into a "grouping problem" and a "shape problem" (cf. Ben-Yosef & Ben-Shahar, 2012b). While the shape problem faces the task of reconstructing a particular contour shape between a given pair of inducers to facilitate recognition and other higher level visual tasks (e.g., Ben-Shahar & Ben-Yosef, 2015; Fantoni, Bertamini, & Gerbino, 2005; Singh, 2004), the grouping problem deals with the process of organizing a set of inducers into pairs between which completion occurs (e.g., Ben-Shahar & Ben-Yosef, 2015; Ben-Yosef & Ben-Shahar, 2012b; Fulvio, Singh, & Maloney, 2008; Jacobs, 1988; Kellman & Shipley, 1991; Shipley & Kellman, 1992). This paper focuses on the latter problem and the visual cues that affect it.

As is well established for more than a century, perceptual grouping of course acts on a scope much broader than visual completion. In general, it is the unconscious process that allows observers to perceive some elements in a visual scene as "belonging together" and, by that, to create semantic precursors for interpretation of the scene (Palmer, 1999; Pomerantz & Kubovy, 1986; Wagemans et al., 2012; Witkin & Tenenbaum, 1983). During the first decades of the 20th century, perceptual grouping has been widely explored by the Gestalt psychologists who suggested a set of qualitative principles, such as proximity, similarity, and good continuation (Pomerantz & Kubovy, 1986; Wertheimer, 1923), that the human visual system employs while processing the basic visual elements of scenes. Although such principles allow valuable insights for predicting the perceptual outcome, Gestalt psychology did not provide quantitative models of how the different perceptual "laws" interact when two or more take effect in a scene. Such quantitative exploration started to emerge only in the past three decades (e.g., Elder & Goldberg, 2002; Feldman et al., 2013; Froyen, Feldman, & Singh, 2010; Geiger, Kumaran, & Parida, 1996; Lugo, Schmiedeler, Batill, & Carlson, 2015), a trend that has not ignored the study of curve completion, as the division into the shape and grouping problems has facilitated further progress.

Indeed, today we are witnessing research on curve completion using a wide range of methods: computational (e.g., Ben-Yosef & Ben-Shahar, 2012b; Hladky & Pauls, 2010; Ming, Li, & He, 2016), behavioral–perceptual (e.g., Maertens & Shapley, 2008; Nayar, Franchak, Adolph, & Kiorpes, 2015; Spelke, 1990), neurobiological (e.g., Ward & Chun, 2016), and inherently multimethodological approaches (e.g., Ben-Yosef & Ben-Shahar, 2012a; Summerfield & De Lange, 2014). Even so, studies on the *grouping* problem have been essentially restricted to the question of what the geometric properties of the occluded and the occluding objects are that determine if a contour will or will not be completed perceptually. More specifically, work on the grouping problem has been restricted to exploring the conditions that allow a *given* pair of inducers to group or not. Perhaps the most influential theory in this context is relatability theory (Kellman & Shipley, 1991; Shipley & Kellman, 1992), which is based on the good continuation principle. This theory suggests that a contour will be completed if inducers are *relatable*, i.e., if the linear extensions of these inducers intersect and the angle between them is right or obtuse. According to the theory, intersection at an acute angle makes the inducer nonrelatable, which prevents grouping and thus results in no completion (see Figure 1).

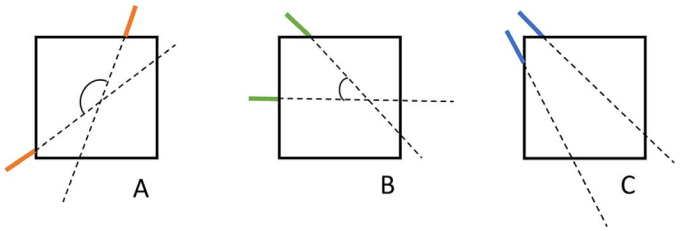

Figure 1. Relatability theory essentials (reproduced from Shipley & Kellman, 1992). (A) Relatable edges. The linear extensions of the red inducers intersect and create an obtuse angle. (B) Nonrelatable edges. The linear extensions of the green inducers intersect and create an acute angle. (C) Nonrelatable edges. The linear extensions of the blue inducers do not intersect at all.

Following relatability theory, in this study we attempt to extend the quantitative exploration of the grouping problem and focus on the wider scope of *multi*-inducer scenarios. Our main goal is studying the factors that effect (and affect) the grouping of *specific* inducer pairs out of all possible inducer groupings in a scene. In particular, we examine the effect of three primary geometric factors that emerge from the Gestalt grouping principles of *proximity* and *good continuation* and were widely considered in the literature in other contexts: relative inducer proximity in the scene (i.e., the relative distances between different inducer pairs), relative inducer orientation, and inducer curvatures.

Methodology

We conducted a perceptual psychophysical experiment to probe how the selected geometric factors influence participants in perceiving different types of completions in a simple multi-inducer completion visual scene. Using the type of perceived completion, we were able to determine the specific grouping of inducers made by the visual system and thus how the geometric properties affected this decision.

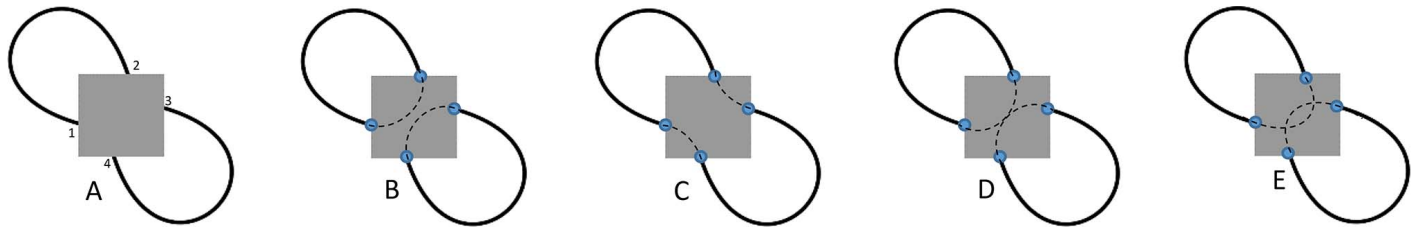

Figure 2. The type of completion based on the inducer grouping. (A) Scene and inducers, the latter numbered 1–4. (B) Grouping of inducers 1, 2 and 3, 4 creates a completion of two separate shapes. (C) Grouping of inducers 1, 4 and 2, 3 creates a completion of a dumbbell. (D) Grouping of inducers 1, 3 and 2, 4 creates a completion of a figure eight. Note that this type of completion is accidental, both geometrically and perceptually. (E) Theoretically, grouping of inducers 1, 2 and 3, 4 as in panel B may also create a completion of two overlapping shapes. However, the last two completion possibilities were excluded in a pilot experiment and thus ignored in our main experiment.

Participants

Twelve students (10 male, two female, all from the department of computer science at Ben Gurion University of the Negev) participated in the experiment for course credit. Two participants were excluded from the analysis due to particularly adversarial performance in experimental repetitions (correlation of 0.51 and 0.52 in all repetitions, respectively). None of the participants were aware of the purpose of the experiment, all had normal or corrected-to-normal vision, and all were right-handed. The experiment was approved by the university ethics committee ahead of time.

Stimuli

For our controlled experiment, we created a simple visual scenario whose center is occluded by a gray square as shown in Figure 2A. This scenario makes four inducers and allows for three possible grouping types and four different completion types: two disconnected convex shapes (Figure 2B), a dumbbell shape (Figure 2C), a figure eight (two shapes that intersect at one point; Figure 2D), or two overlapping shapes (Figure 2E). The figure eight–type of organization is accidental (both geometrically and perceptually) and ambiguous in terms of grouping, so in order to understand if and how participants perceive this interpretation, or the ambiguous overlapping one, we first conducted a pilot experiment in which we presented a preliminary (and separate) group of observers stimuli similar to those used later in the main experiment but asked them to report the grouping of inducer pairs as perceived. After the pilot experiment itself, we also debriefed the subjects about the object or objects they perceived. No (i.e., zero) subject reported the perception of two overlapping objects as in Figure 2E or the figure eight organization as in Figure 2D. Hence, for the study reported in this paper, we focused only on the first two configurations as shown in Figure 2B and C and thus on resolving whether participants grouped the inducers to create one or two shapes.

As mentioned above, in this research we focused on three independent geometrical factors of the inducers and explored how they interact to affect the grouping and completion results. In order to accurately manipulate the curve's properties at the occlusion points but also obtain a closed and compact shape, we first created the inducing segments next to the occlusion point and then linked these segments with a smooth interpolation (see below). In all cases, we discretized the relevant range of each factor to five discrete values in order to end up with a manageable number of trials.

Driven by the most fundamental Gestalt principle (Wertheimer, 1923), our first manipulated factor is the relative proximity, or the proximity ratio (r) between the inducers, i.e., the ratio of distances (in millimeters or visual angle) between the relevant pairs of inducers (see Figure 3A). The measured levels varied between equal distances between the two pairs (21.2 mm on screen, visual angle of 1.215° from the designated viewing distance, $r = 1$), one pair of inducers is moderately closer than the second pair (the closer pair is 19.37 mm or 1.11° apart, and the other pair is 23.05 mm or 1.32° apart; $r = 0.84$ and $r = 1.2$ in this case, respectively), and one pair is significantly closer than the other (the closer pair is 17.67 mm or 1.012° apart, and the other pair is 24.74 mm or 1.417° apart; $r = 0.71$ and $r = 1.498$ in this case, respectively).

Following relatability theory (Kellman & Shipley, 1991), our second manipulated factor was the orientation of the inducers at the occlusion point or, more accurately, the relative orientation of the different inducer pairs as measured by the angle between the linear extensions of their tangents at the occlusion point. This angle could be acute, right, or obtuse between each pair of inducers and its complementary angle applied to the other two tangents (see Figure 3B). Here our stimuli allowed a right angle (90°), one acute angle (77.8°), and three obtuse angles (101.8°, 113.2°, 124.3°). The distribution of tested angles is not

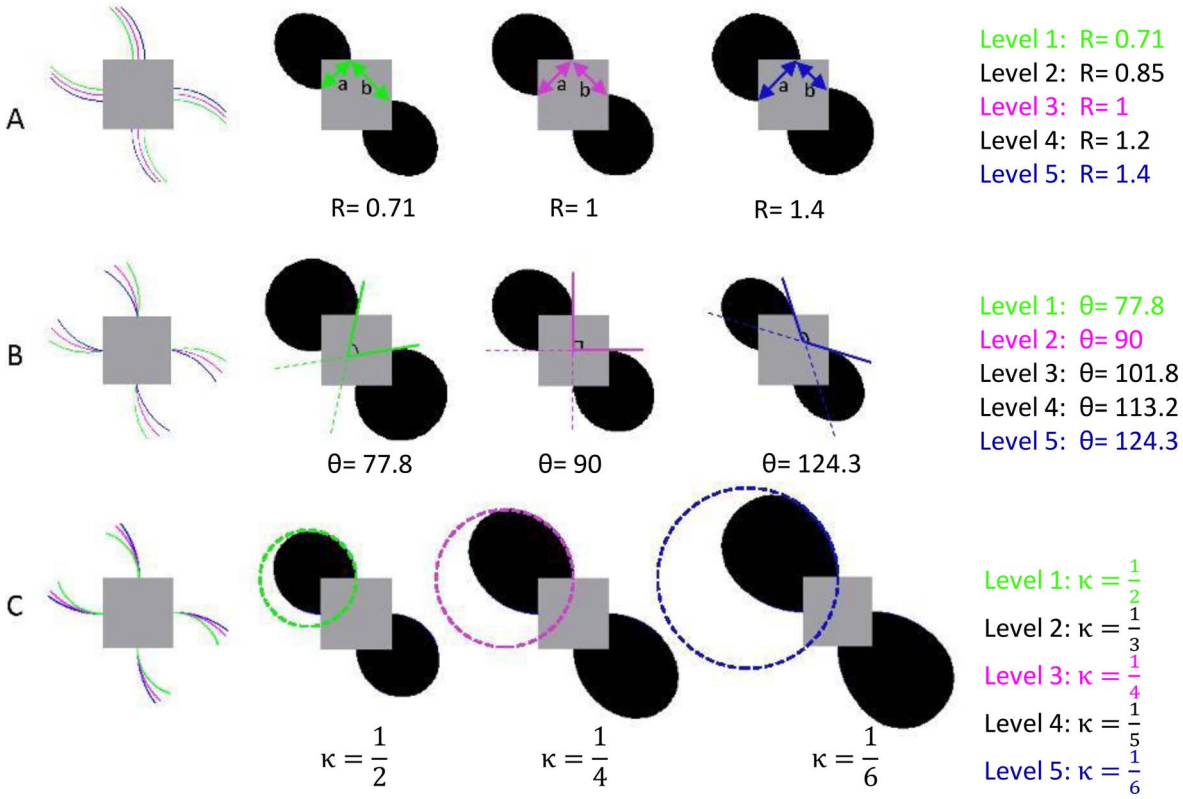

Figure 3. Illustration of the values of the different geometrical factors of the inducers at the occlusion points and examples of the occluded shapes they entail. To avoid clutter, in all cases we show just three out of the five values tested. (A) The first independent stimulus factor is the "proximity ratio" measured as a/b, i.e., the ratio of the distances between one inducer pair to the other one. (B) The second stimulus factor is the angle between the inducer tangents. This angle can be acute, right, or obtuse. (C) The third factor is the curvature of the occluded curve at the inducer. To facilitate a stable curvature value, both perceptually and computationally, the contour extended from the inducer is a circular arc of radius 1/κ and predetermined and fixed arc length. Superimposed on the selected examples here are also the full circles of which these arcs are members.

symmetric about 90° because preliminary results from a second pilot study (see below) found a ceiling effect in subject responses for angles smaller than 77.8° or larger than 124.3°, a range we later divided into roughly equal intervals.

The last manipulated factor was inducer curvature (do Carmo, 1976), i.e., the curvature of the occluded curve at the occlusion point, a property known to affect the shape problem in curve completion (Ben-Shahar & Ben-Yosef, 2015; Takeichi, Nakazawa, Murakami, & Shimojo, 1995) but also other visual tasks, such as curve inference (e.g., Field, Hayes, & Hess, 1993; Parent & Zucker, 1989) and orientation-based texture segregation (e.g., Ben-Shahar & Zucker, 2004; Bhatt, Carpenter, & Grossberg, 2007), to name but a few. In order to obtain a stable curvature value (both geometrically and perceptually) at the occlusion point, we made the first segment of the observable part of the contour a circular arc, whose curvature is fixed and equal to the inverse of the radius (i.e., $k = \frac{1}{R}$). Hence, in order to manipulate curvature, we modified the radius of the circular segment while keeping its arc length constant across stimuli. To produce arc length of L length units (say, centimeters) on screen, a circular arc of radius R should extend $\frac{L}{R} = Lk$ radians in angle. In our case, the tested curvature values were set to be $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$, $\frac{1}{5}$, and $\frac{1}{6}$, and the fixed arc length was selected to be 2.5 cm (on screen). This resulted in circular arcs of $\frac{2.5}{2}$, $\frac{2.5}{3}$, $\frac{2.5}{4}$, $\frac{2.5}{5}$, $\frac{2.5}{6}$ radians in angle, respectively. The effect of this factor on the overall observable shape is illustrated in Figure 3C.

The parameter range for all three parameters in the main experiment was selected in a second pilot experiment in which subjects were asked to complete a two-alternative forced choice (2AFC) task and provide an explicit yet qualitative response on whether they perceived two blobby shapes or one dumbbell shape. In addition to the paradigm used, the pilot experiment also differed from the main experiment in using a wider range of parameters. The ceiling conditions for the main experiment were determined by the stimulus conditions that yielded *identical* responses across 11 pilot subjects across repetitions. The pilot trials were omitted from analysis of the main experiment, but their

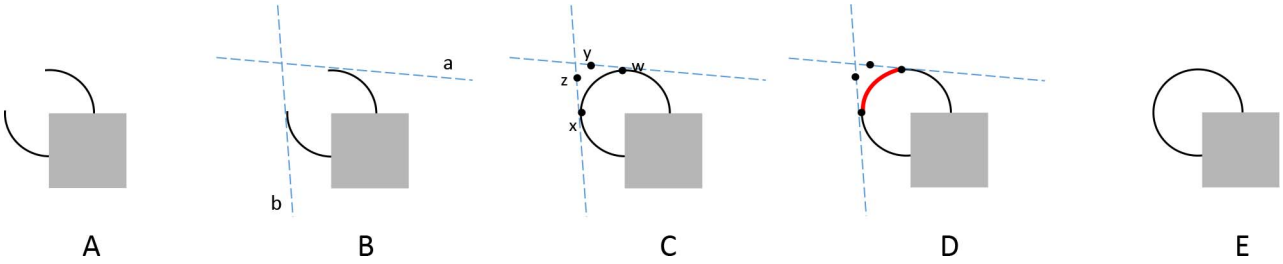

Figure 4. Generating the occluded contour from circular arcs and Bézier interpolation. (A) The basic inducing curves are created as circular arcs at the occlusion points after setting their properties (distance ratio, orientation, and curvature). (B) The tangent lines a and b are extended from the far end points of the circular arcs. (C) The initial points of the Bézier interpolant are set at the end points of the circular arcs (points w, x), and the control points are set further away on the tangent lines (points y, z). (D) A smooth Bézier curve (in red) is interpolated using the control points. (E) The full closed contour that delimits the shape presented to the observer (as in Figure 3).

ceiling conditions were used as stimuli during its training phase (see below).

As mentioned above, our shapes were created from inducing circular segments (with known position, orientation, and curvature) at the occlusion points, linked together via smooth interpolation—in our case a Bézier interpolation (Figure 4). To do so, we set the far end points of the circular segments (**w,x** in Figure 4C) as the initial points of the Bézier segments and added two additional points (**y,z** in Figure 4C) on their linear extensions (**a,b** in Figure 4). To avoid shape artifacts and end up with roughly similar shapes in all cases, these additional points were positioned at a distance influenced by the inducer curvature, and in particular, as the curvature got smaller (i.e., the radius of the circle got bigger) the additional control points were located farther away from the circular arcs. The entire process is illustrated step by step in Figure 4.

Procedure

In order to minimize task-related confounds or biases, we did not ask participants to complete an explicit grouping or completion tasks. Rather, we sought these visual decisions only implicitly by employing a dot localization paradigm (Guttman & Kellman, 2004; Pomerantz, Goldberg, Golder, & Tetewsky, 1981; Unuma, Hasegawa, & Kellman, 2010). In this paradigm, the participant is required to make a 2AFC whether a presented dot probe is *inside* or *outside* a query shape. In our case, subjects were asked to "decide whether the flashed probe dot was located inside or outside the occluded object or objects," and the latter phrase was used to eliminate bias and imply that there could be one completed object or more. Indeed, implicit measures might be ambiguous in their mapping to the desired dependent parameter. But, in our case, the automatic character of dot localization allowed for rapid stimulation and exploration of bottom-up perceptual processes while minimizing the involvement of top-down or higher cognitive processes that might otherwise occur during explicit measures (such as when asking subjects to plot the completion type or the grouping relationship).

In the experiment, for each stimulus condition, we flashed such a probe dot in one of five different locations. Dots were distributed uniformly along the diagonal of the occluding square while alternating sides off its center. Positioned −0.4 cm, −0.2 cm, 0 cm, 0.1 cm, and 0.3 cm relative to the center of the square, the eccentricity of these probe dots was −0.229°, −0.115°, 0°, 0.057°, and 0.172° of visual angle, respectively. The dots were all identical in size (2.47 mm on screen or approximately 0.14° of visual angle from the designate viewing distance). Participants were required to respond whether each dot was "inside" or "outside" the occluded shape/shapes. The presentation order of probed locations was random. By combining the in/out answers for all five probes, we were able to determine the participant's implicit completion type and, consequently, the specific grouping of inducers, i.e., grouping of inducers 1, 2 and 3, 4 or 2, 3 and 1, 4 (see Figure 5).

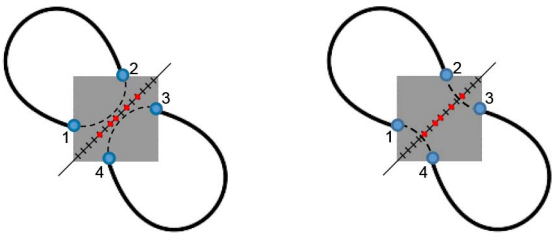

Figure 5. Inferring grouping decisions from dot localization. By combining the different trials for each inducer condition, we were able to determine whether the participant perceived two shapes or one and thereby whether the participant grouped 1, 2 and 3, 4 (none of the probes were in) or perhaps 2, 3 and 1, 4 (some of the probes were in). The completed (dashed) arcs are merely schematic and do not necessarily reflect a perceived shape in this stimulus.

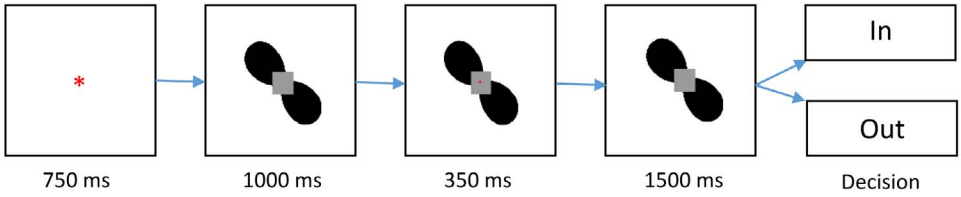

Figure 6. The experimental procedure (left to right): A fixation point was followed by a 1000-ms stimulus presentation. The probe dot was then flashed for 350 ms, after which the subject had up to an additional 1500 ms to respond (while this stimulus was still presented but without the probe dot).

Each participant was thus presented with 125 stimulus conditions (5 proximity ratios × 5 relative angles × 5 inducer curvatures) × 5 probe locations for a total of 625 trials in a session. To increase the power of results per stimulus condition, each observer participated in two such sessions (in two consecutive days) to double the number of trials.

The experiment began with a short message that described the procedure. In each trial, the participant first needed to fixate on a fixation point for 750 ms. The stimulus, with the center of the occluding square overlapping the fixation point, was then presented for 1000 ms, after which the red probe dot flashed for 350 ms. The participant then had up to an additional 1500 ms to view the object without the probe and to provide an "inside" or "outside" response by pressing one of two designated keys (see Figure 6). The experimental design was completely unblocked and randomized. Subjects were offered four breaks (thereby splitting the experiment into five sessions) during which they were asked to let their eyes rest.

Before the experiment started, the participants completed a practice session of 10 trials. The main goal was to verify their understanding of the instructions, and thus, they were presented with ceiling conditions only (i.e., stimuli conditions that were not used in the actual experiment and were far from the decision threshold as found in the pilot study mentioned above) and received correct/incorrect feedback after their responses.

Apparatus

All stimuli were programmed in MATLAB (2014b), and the behavioral procedure was implemented in E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA). The experiment was performed in a small, dimly illuminated room that was designed especially for behavioral studies. Participants were seated approximately 100 cm from the screen with their two index fingers on the response keys. Experiment hardware constituted a 3GHz Intel core 2 Duo processor and a 22-in. Samsung 2243BW LCD screen at 1,680 × 1,050 resolution and 60-Hz refresh rate.

Analysis

After the behavioral experiment, we determined the completion type (and therefore the grouping type) of each stimulus condition from the response pattern of each participant. For each of the 125 different stimulus conditions, we combined the in/out responses for the five probes to determine whether the stimulus was perceived as one object or two objects. In particular, a participant who responded "out" for all five dot probes in a specific stimulus indicated the perception of two different shapes. A participant who responded "in" for the center probe (and possibly other probes also) indicated the perception of one shape. Finally, the trials in which a participant responded "out" for the center probe but "in" for at least one other probe were marked "inconsistent" and excluded from the analysis.

We assigned the two completion types the values zero (completion into one shape) and one (completion into two separate shapes) and then averaged across subjects to obtain the probability of a two-shape completion (and thus the one for one-shape completion also). For example, the stimulus with $r = 0.84$, $k = \frac{1}{5}$ and $\theta = 90°$ was perceived by four participants as one shape (answered "in" for at least the center probe, assigned the value zero) and by eight participants as two shapes (answered "out" for all the dots, assigned the value one). In this case, the average of participant responses thus indicates 0.66 probability to perceive two shapes and 0.33 probability to perceive one.

Because we have three independent parameters, the probability to perceive one or two shapes can be assigned to a three-dimensional space in which one can perform queries, find a decision surface, or obtain other insights. Because visualizing a 3-D function is difficult using 2-D figures, we present the results as a collection of cross-sections along the different axes of the space. Such a presentation essentially fixes one parameter while varying the other two, an operation that can also reveal relationships between the three parameters. To permit dense visualization, the discrete samples are converted to a continuous function in the parameter space by linear interpolation.

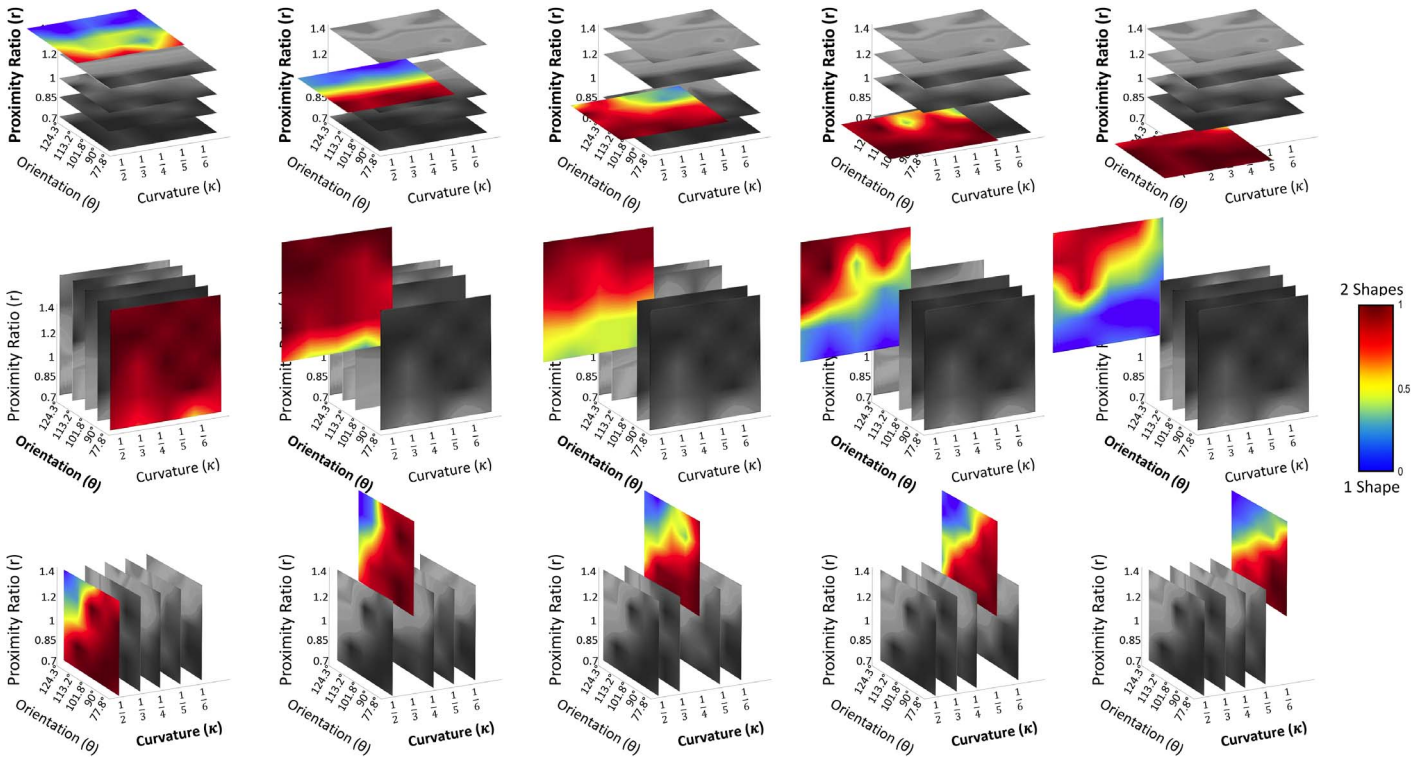

Figure 7. Visualization of the three-dimensional perceptual function as a function of the three explored stimulus factors. The axes refer to our independent parameters, and each row focuses on cross-sections by one parameter. Red regions of the parameter space refer to higher probability to group inducers 1, 2 and 3, 4 and perceive two shapes. Blue regions refer to parameter combinations that entail higher probability for grouping inducers 1, 4 and 2, 3 and perceiving one shape.

Results

As discussed above, we collected data about the probability to perceive a specific completion type (thus a particular grouping configuration) for each and every combination of stimulus parameters. Fourteen point seven percent of all trials (across subjects, conditions, and repetitions) were found "inconsistent" (see the Analysis subsection) and excluded from the analysis. The results of all valid trials are summarized in Figure 7 by presenting all cardinal cross-sections of the three-dimensional response function in the independent parameters space. The following are general observations and properties of this response function.

Scrutinizing the results in Figure 7 along the proximity ratio axis, it is clear that the smaller the proximity ratio, the higher the probability to perceive two shapes (i.e., grouping of inducers 1, 2 and 3, 4) rather than one as is evident from the increasing portion of red area once we move from left to right in the top row in Figure 7. Clearly, when the proximity ratio is large (e.g., $r = 1.2$ as shown in the close up in Figure 8A), the probability to perceive one shape is not only significant, but it also depends on other factors and, in particular, on the relative orientation. On the other hand, when the proximity ratio is small,

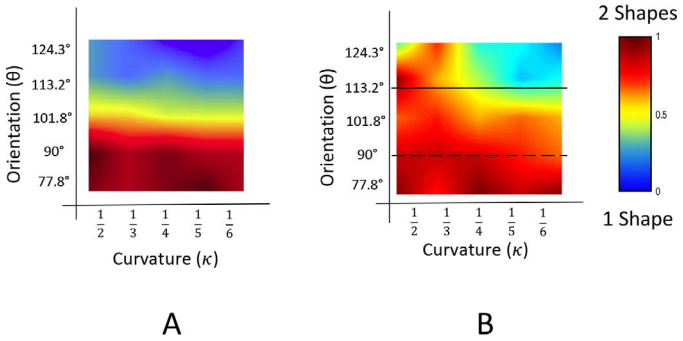

Figure 8. (A) The perceptual result for proximity ratio $r = 1.2$. Note how it is strongly affected by the relative orientation between the different pairs of inducers and how this factor alone can change the grouping result as we move top to bottom on the cross-section. Note that because the level sets of the response function in this case are parallel to the curvature axis, inducer curvature plays little role in the perceptual outcome in the particular case. (B) The perceptual result for proximity ratio $r = 1$. Note first the bias for the perception of two shapes (red color) when the proximity ratio and the relative orientation are both at equilibrium ($r = 1$ and $\theta = 90°$, dashed line). In other cases (e.g., $\theta = 113.2°$, solid line), the effect of curvature is pronounced as changing only that value can completely reverse the perceptual result (as evident by the changing color along that solid line).

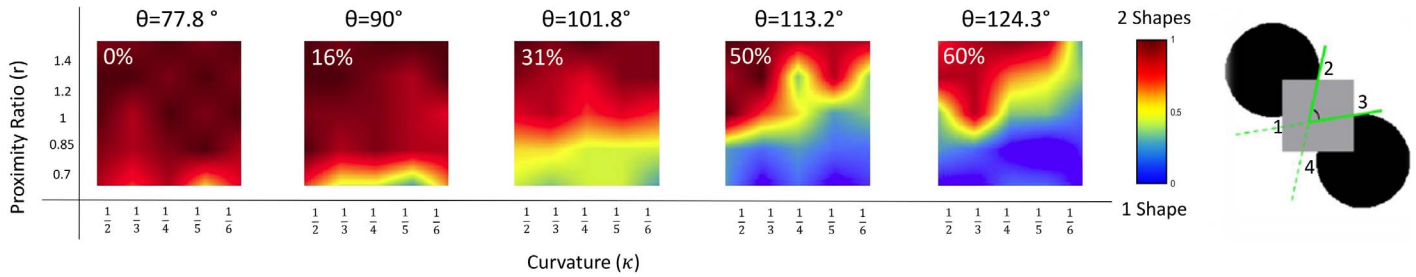

Figure 9. The probability of perceiving one shape rather than two for each relative orientation. The numbers on the top left corners denote the total probability of perceiving one shape rather than two for that relative orientation condition (when pooled over all proximity ratios and curvatures).

performance saturates at two shapes regardless of the other factors. That said, it is particularly interesting to examine the $r = 1$ cross-section (top middle panel in the first row of Figure 7 and Figure 8B), i.e., the condition in which proximity is essentially factored out from the perceptual outcome (because it provides no bias toward any specific pair of inducers). Perhaps expectedly, the result is not fixed, and in particular, it is not independent of the relative orientation between the inducers. But perhaps less expectedly so, when the relative orientation is set at the physical equilibrium of 90° (dashed line in Figure 8B), the result is clearly biased at *two* shapes (as opposed to a chance level performance as entailed by relatability theory). Moreover, for larger relative orientation between the inducer pairs, the perceptual result clearly depends on inducer *curvature* (e.g., note how the color of the map changes as we sweep along the solid line at $\theta = 113.2°$ in Figure 8B).

The second geometric parameter we examined is the relative orientation between inducers (middle row in Figure 7). Recall that the independent parameter θ reflects the orientation between inducers 2 and 3 (see Figure 3), and the complimentary angle $(180° - \theta)$ holds between inducers 1 and 2. In agreement with relatability theory (Kellman & Shipley, 1991), we found that when θ is acute (77.8°, left panel, middle row in Figure 7), participants refrain from grouping them together and thus perceive two separate shapes for all curvature–proximity ratio combinations. This tendency declined gracefully as the relative orientation increased. Indeed, when pooled over all subjects and all other stimulus factors (i.e., all possible proximity ratios and curvatures) the probability to group inducers 2, 3 and 1, 4 (and thus perceive one shape) was 16%, 31%, 50%, 60% for $\theta = 90°$, 101.8°, 113.2°, and 124.3°, respectively (see Figure 9). What is particularly interesting, however, is how the perceptual function becomes less homogeneous and how its level sets become less parallel to the curvature and proximity ratio axes as the relative orientation becomes larger. In other words, the grouping decisions became progressively more influenced by both proximity ratio and inducer curvature as relative orientation increased.

Finally, the effect of curvature on the results is observed explicitly in the last row of Figure 7 and in Figure 10, where we find that as curvature decreases, so does the probability of perceiving two shapes rather than one (75%, 71%, 67%, 64%, and 61% for curvature values $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}$, respectively, when pooled over all proximity ratios, all relative orientations, and all subjects). Clearly, the perceptual result for some combinations of orientation and proximity ratios saturates at two shapes and is not modulated by curvature (e.g., note how all observers perceived two shapes for $\theta \leq 90°$ and $r \leq 0.85$ regardless of inducer curvature). However, curvature has a categorical effect closer to the perceptual threshold and could completely turn the grouping result upside-down to switch perception from two objects to one (see marked coordinates in Figure 10).

Modeling

Although Figure 7 can be considered a model of the perceptual function that governs inducer grouping in multi-inducer stimuli (such as the one used in our experiment), it provides little insight regarding the relationship between the independent variables and little explanation as to how they interact in dictating the perceptual outcome. To achieve this goal, a parametric model is much more constructive. Although such a full-fledged predictive model is beyond the scope of our present work, in this section, we show that much of the behavior represented by the map in Figure 7, and a surprising interaction between two of the geometrical factors we explored, can be explained by the simplest type of formalization.

Qualitatively speaking, the response map in Figure 7 exhibits level sets that in many panels appear approximately linear although not necessarily parallel to any of the axes. Attempting to formalize this observation and in order to explore such a possible relationship

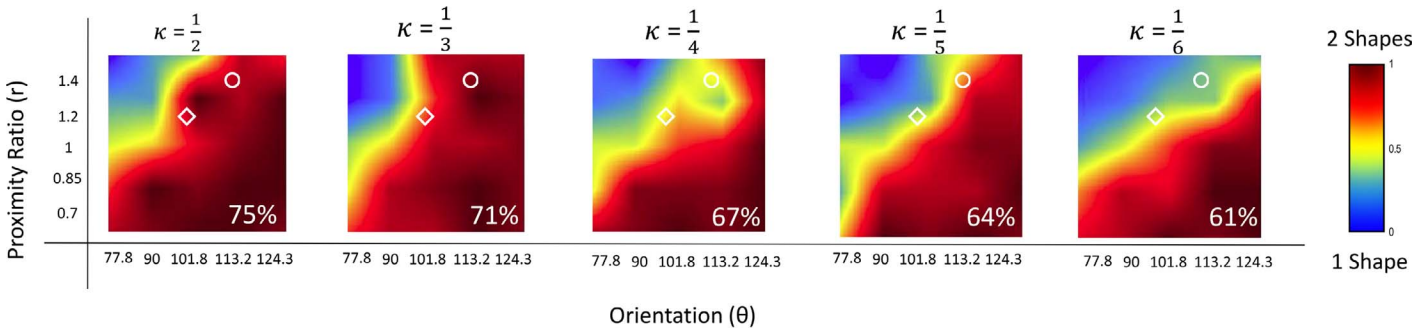

Figure 10. The probability of perceiving one shape rather than two for each curvature level. The numbers on the bottom right corners denote the total probability of perceiving two shapes rather than one for that curvature condition (i.e., pooled over all proximity ratios and relative orientations). Note how for some combinations of proximity ratio and relative orientation the response function changes categorically with curvature alone. (e.g., note how the color inside the circles for when $\theta = 113.2°$ and $r = 1.4$ or inside the diamonds for $\theta = 101.8°$ and $r = 1.2$ is not identical in all panels).

between proximity ratio, relative orientation, and curvature as they are mapped to the perceptual grouping result, we considered the following rudimentary linear model:

$$P = P[\text{observer perceives two shapes}] = P(r, \theta, \kappa) = \alpha_0 + \alpha_r \cdot r + \alpha_\theta \cdot \theta + \alpha_\kappa \cdot \kappa \quad (1)$$

where (as defined in Figure 3) r is the proximity ratio between the inducers; θ is their relative orientation; κ is their curvature; α_0 is an inherent bias to perceive two shapes as opposed to one; and α_k, α_r, and α_θ weigh the relative contribution of each of the three parameters in affecting the final perceptual result. To find the values of these weight coefficients that closely approximate the behavioral function P observed in our experiment, we use standard least squares optimization methods. Because the different parameters are not commensurable (neither in units nor in the range of their values), this process was applied on their index values rather than their real values, on which the proper mapping is deferred for later (see below). The result obtained is the following explicit model:

$$P[\text{perceive two shapes}] = 0.76 - 0.033 \cdot \kappa + 0.14 \cdot r - 0.14 \cdot \theta \quad (2)$$

This simple model fits the data with a root-mean-square error (RMSE) of 0.11, and if a decision threshold is set at 0.5 (in order to produce a forced-choice response), it agrees with 116 of 125 recorded observations, i.e., it performs at 92.8% accuracy.

Scrutinizing the model coefficients reveals several interesting findings. First, $\alpha_0 > 0.5$ implies an inherent bias of the visual system to a grouping result that entails *two* shapes. Second, curvature contributes less (α_k is smaller than its counterparts) but clearly does contribute to the perceptual outcome. Finally, and perhaps most importantly, proximity and orientation seem to exhibit an opponency relationship (because $\alpha_r = -\alpha_\theta$). In order to further examine this relationship, we applied principal component analysis to the observed data (i.e., both the independent and dependent parameters) and obtained the principle components (PCs) detailed in Table 1.

Not only is the linear manifold of observations three-dimensional (note that three PCs are enough to describe almost all variance), implying that all three factors affect the distribution of responses, but all PCs have equal length projection on the r and θ axes (the sign in this case is irrelevant), implying equal contribution of those dimensions to the representation of the linear manifold of the observed data. Furthermore, all PCs with a nonvanishing P coordinate (i.e., PC_1 and PC_4, the only one who can contribute to a nonvanishing value of P in the eigenspace) exhibit not only magnitude equality but also opponency behavior (i.e., opposite sign). We conclude that opponency between proximity ratio and orientation plays a basic role in inducer grouping.

To double check the insights above, we again sought to model the data, but this time with a simpler three-parameter model in which this opponency is represented intrinsically, i.e.,

$$P = \alpha_0 + \alpha_{r\theta} \cdot (\theta - r) + \alpha_\kappa \cdot \kappa \quad (3)$$

Least squares optimization methods were again applied, resulting in similar coefficients and accuracy as those of the unconstrained model (RMSE 0.11, 92.8% accuracy):

Principle component	PC1	PC2	PC3	PC4
r	0.69	0.13	0.70	−0.13
θ	−0.68	−0.13	0.71	0.13
κ	−0.18	0.98	0.0	0.034
P	0.19	0.0	0.0	0.98
Variance explained, %	34%	33%	33%	≈0

Table 1. Principle components (PCs) obtained by applying principal component analysis to observation data. Note that PCs with a nonvanishing value of P (PC1, PC4) point toward an opponency between proximity ratio and orientation.

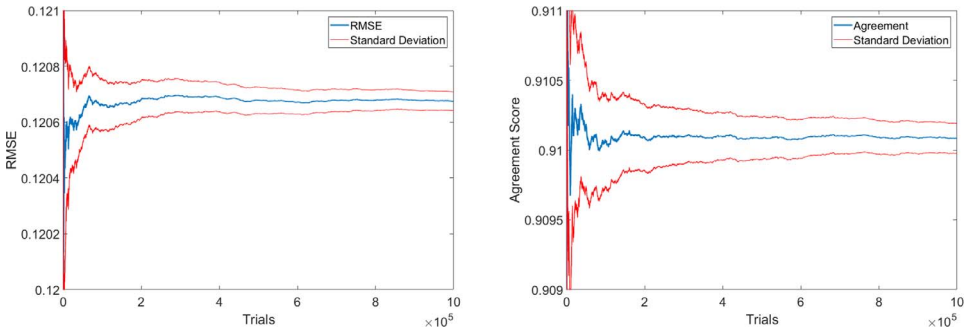

Figure 11. Running average (blue) of both RMSE (left) and prediction accuracy (right) and the corresponding standard deviations (red envelope) of the linear model as a function of number of trials (from one through 10^6). In each trial, the model was trained from a random half of the experimental data and tested on the other half.

$$P[\text{perceive two shapes}] = 0.76 + 0.14 \cdot (r - \theta) - 0.033 \cdot \kappa \quad (4)$$

Figure 12 presents this model's predictions over the entire parameter space, similarly to Figure 7.

Although it is not necessarily the case that fitting a three-parameter model to the observed data should necessarily produce such a good fit as above, we nevertheless explored the robustness of this model by splitting the data to "train" or "test" sets with a probability of 50%. The constrained model was trained exclusively over the training set and its performance (i.e., predictive power) measured on the test set. This experiment was repeated many times (each time with a different random split to train and test sets) with both average RMSE and average response prediction success rate calculated from the results. Figure 11 shows these results as a function of the number of repetitions. Note that both in individual experiments and on average after many repetitions, both RMSE (Figure 11, left) and prediction accuracy are similar to the model when fitted based on all data (RMSE = 0.12 compared to 0.11 above and prediction accuracy >91% compared to 92.8% above). Note again that this is while considering only half the observations.

Finally, recall that the model obtained above is based on the index or "level" of parameters (as defined in Figure 3). This was necessary in order to make the parameters commensurable despite having different

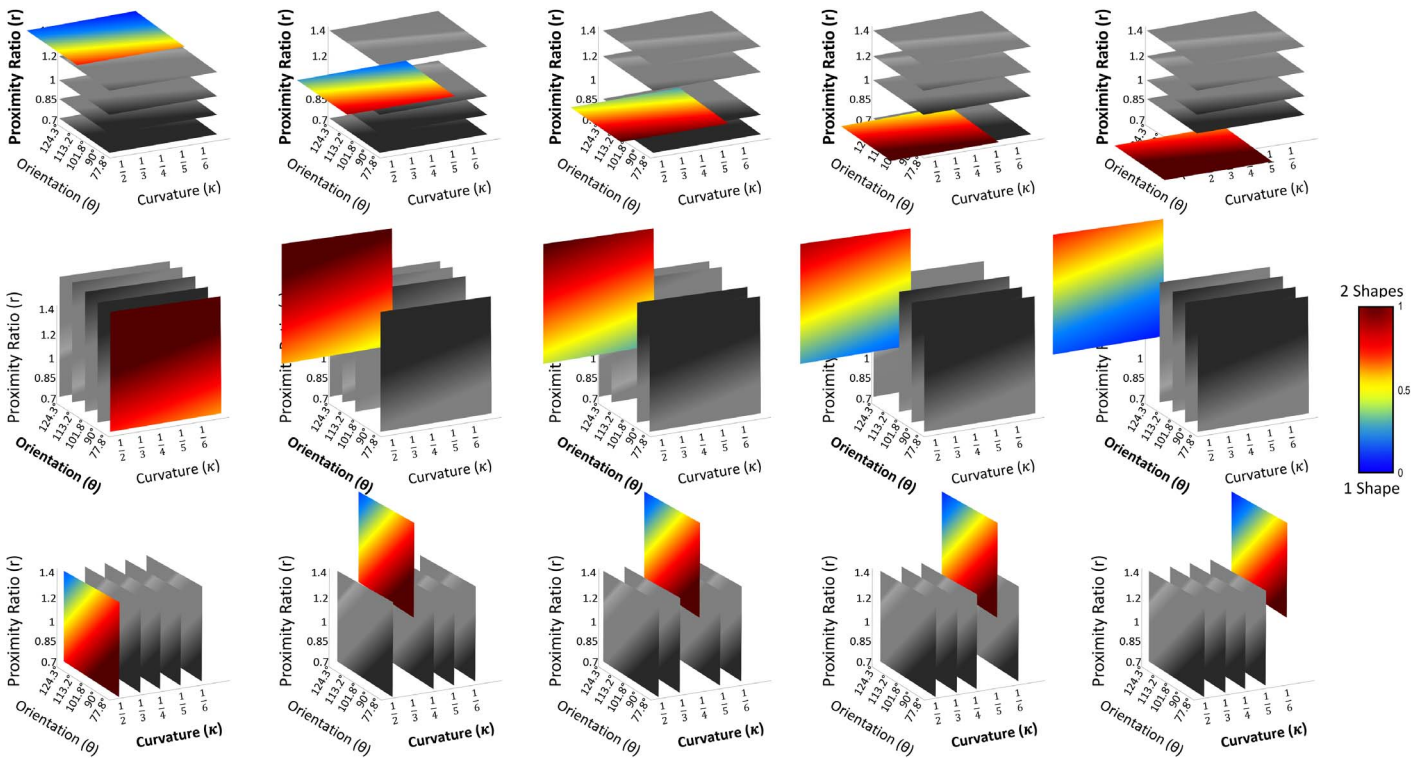

Figure 12. Visualization of the three-dimensional perceptual modeling function, color-coded by probability similar to Figure 7.

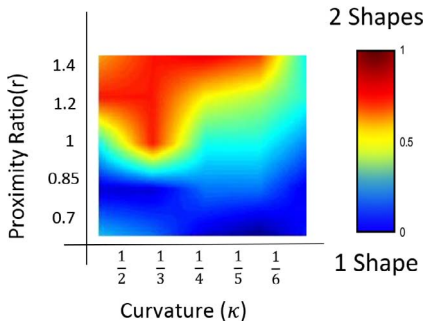

Figure 13. The behavioral results for orientation $\theta = 124.3°$. Although relatability theory would predict the same perceptual outcome in this subset of the parameter space, in practice the latter is strongly modulated both by the proximity and the curvature of the inducers (as reflected by the nonuniform value of the response function over the cross-section).

units and overwhelmingly different ranges and scale of values. However, it is easy to transform this representation to one in which the true/real values of the parameters can be plugged in for prediction of the outcome in novel conditions. This type of conversion and the real-valued model are specified in the Appendix.

In conclusion, and as mentioned in the Introduction, although Gestalt psychologists provided valuable insights for predicting the perceptual outcome in simple stimuli, they did not provide quantitative models of how their different perceptual laws interact when two or more take effect in a scene. The simple initial model that emerges above suggests that, at least in the context of multi-inducer grouping, the interaction between proximity and good continuation (expressed by relative orientation) takes a relatively simple opponency form with particular weights.

Discussion

In this study, we examined how our visual system groups multiple inducers into pairs in order to facilitate completion of missing visual information in ambiguous scenes. To make our study tractable, we focused on (and manipulated) three independent geometrical features: proximity ratio, relative orientation, and curvature of the inducers at the occlusion point. Grouping decisions by subjects were extracted implicitly using a dot localization paradigm and resulted in a 3-D response function that represents how the three independent parameters conspire to produce the perceptual outcome.

The 3-D response function indicates first that proximity plays a (classical) crucial role at the grouping decision and that the latter is strongly affected by the relative orientation of the inducers, i.e., their degree of good continuation (Figure 8A). Although this finding is missing from prior theories, we found that curvature also modulates grouping results, especially when proximity is at (or close to) equilibrium (Figure 8B).

Unlike the above, the effect of good continuation (which in our case was reflected by the relative angle parameter) was expected. Clearly, the more obtuse the angle between the linear extension of the inducers, the more likely they are to group together to induce a completed curve. However, *unlike* the prediction by the classical relatability theory, this tendency is strongly modulated both by proximity and by the curvature of the inducer as shown by the nonuniform response in Figure 13.

The effects of inducer curvature are illuminating in at least two ways. First, the results clearly show that there are conditions that are affected by inducer curvature and inducer curvature alone (see Figure 14).

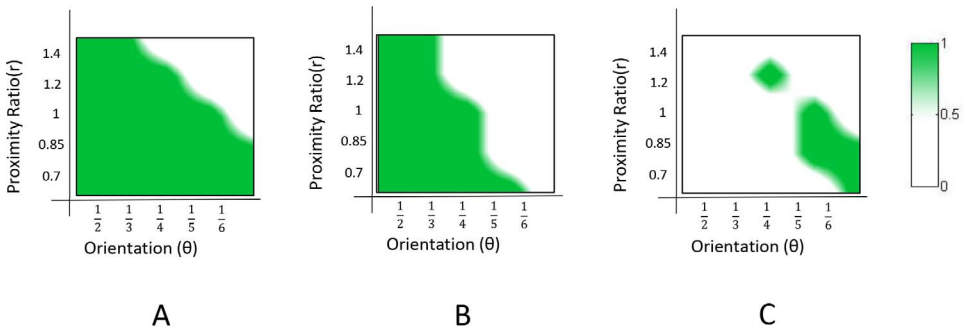

Figure 14. Although confined to a relatively small subset of the parameter space, inducer curvature affects the grouping results. The maps shown here are the likely perceptual response of a subject, computed by rounding the corresponding response panels from Figure 7 to the nearest integer (where zero denotes the perception of a single object and one denotes two objects). (A) The likely response function for the biggest curvature tested ($\kappa = \frac{1}{2}$). (B) The likely response function for the smallest curvature tested ($\kappa = \frac{1}{6}$). (C) The absolute difference map of the two response functions. Any positive value in this map represents a stimulus condition for which the mere change in inducer curvature, while keeping all other stimulus properties fixed, entailed a qualitative categorical change in the perceptual outcome.

Figure 15. An original demo in the context of grouping is the following one: On the left panel, most observers report chance level grouping of the left contour with either of the contours to its right. At the same time, in the two rightmost panels, observers are found to be biased toward a single grouping. Note that inducer positions and orientations are identical in all three cases.

This stands in sharp contrast to existing theories (i.e., relatability) but is consistent with earlier reports on the effect of curvature in the context of the shape problem (e.g., Takeichi, 1995) and the role of curvature in general on the representation and processing of shape (see Figure 15). Second, the effect of curvature appears to be similar to making the angle between (the linear extension of the) inducers more obtuse; larger inducer curvature and/or obtuse inducer angles resulted in a greater probability to complete the scene into two different shapes. Unlike proximity and relative orientation, which appear to inhibit each other (at least under the conditions of our experiment), relative orientation and curvature seem to facilitate each other. Indeed, both of these behaviors are reflected by the sign of the corresponding coefficients in our approximated linear model (Equation 4).

With all of proximity, relative orientation, and curvature now expressed explicitly in our response function and model, it is particularly tempting to compare the behavioral results with the predictions of relatability theory (Kellman & Shipley, 1991) in the entire domain. In fact, one could easily produce the latter in a form similar to Figure 7, and by doing so, we obtain Figure 16. Naturally, the response function of relatability in this 3-D parameter space is constant both along the proximity ratio axis and the curvature axis. But, with this explicit representation, we can now obtain a clear representation of the (absolute value) difference between relatability and the generalization we studied here for multi-inducer scenes. Figure 17 depicts the level of disagreement between the participants' decision and the relatability theory prediction. As shown, a significant portion of the parameter space exhibits a difference greater than 0.5 (a conservative threshold to represent change in the likely perceptual outcome), indicating the limitation of relatability theory and the fundamental need to extend it (both empirically and theoretically) to the parameter space suggested here. Quantitatively, repeating the performance analysis from the Modeling section, but this time applying it to the predictions of relatability theory, we find that it has a much larger RMSE of 0.42 (compared to 0.11 in our proposed model) and a significantly lower prediction power of just 68%

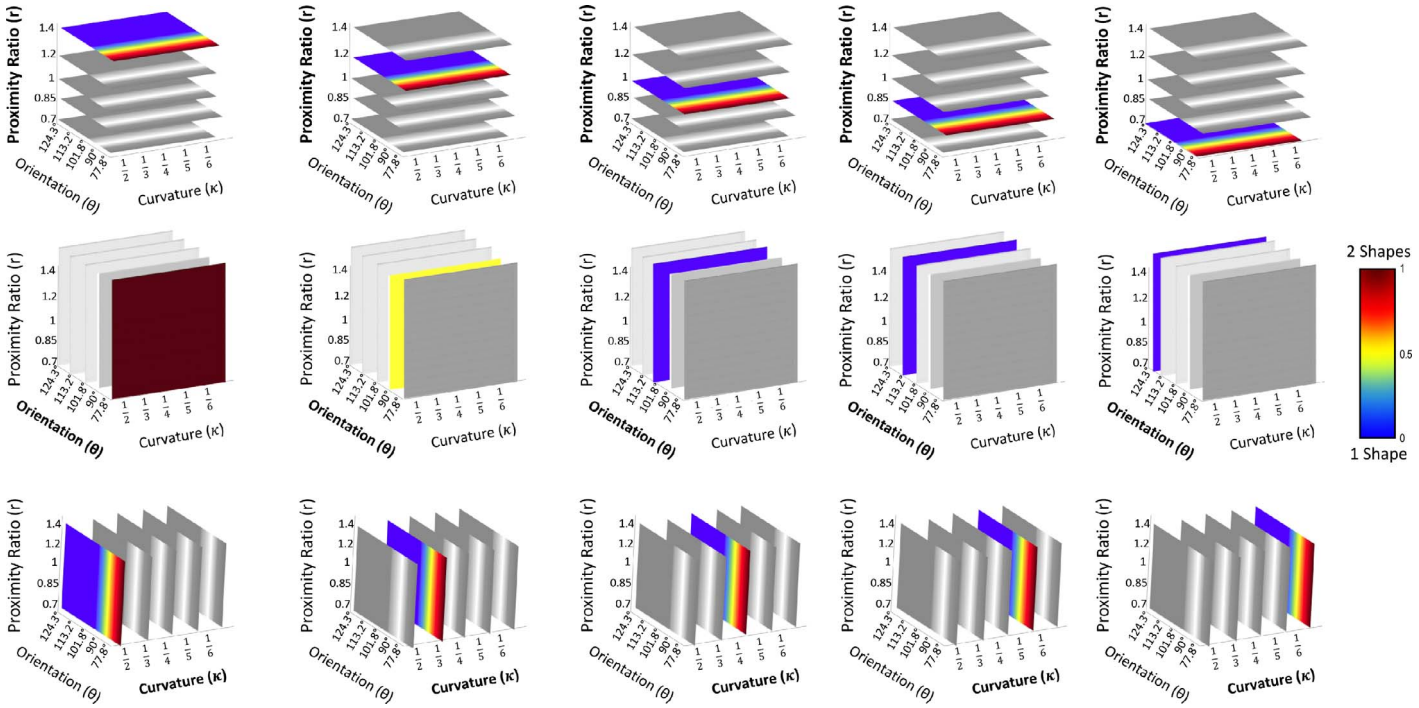

Figure 16. An illustration of the predictions of relatability theory depicted in a way similar to our behavioral results presented in Figure 7 (see also Methodology).

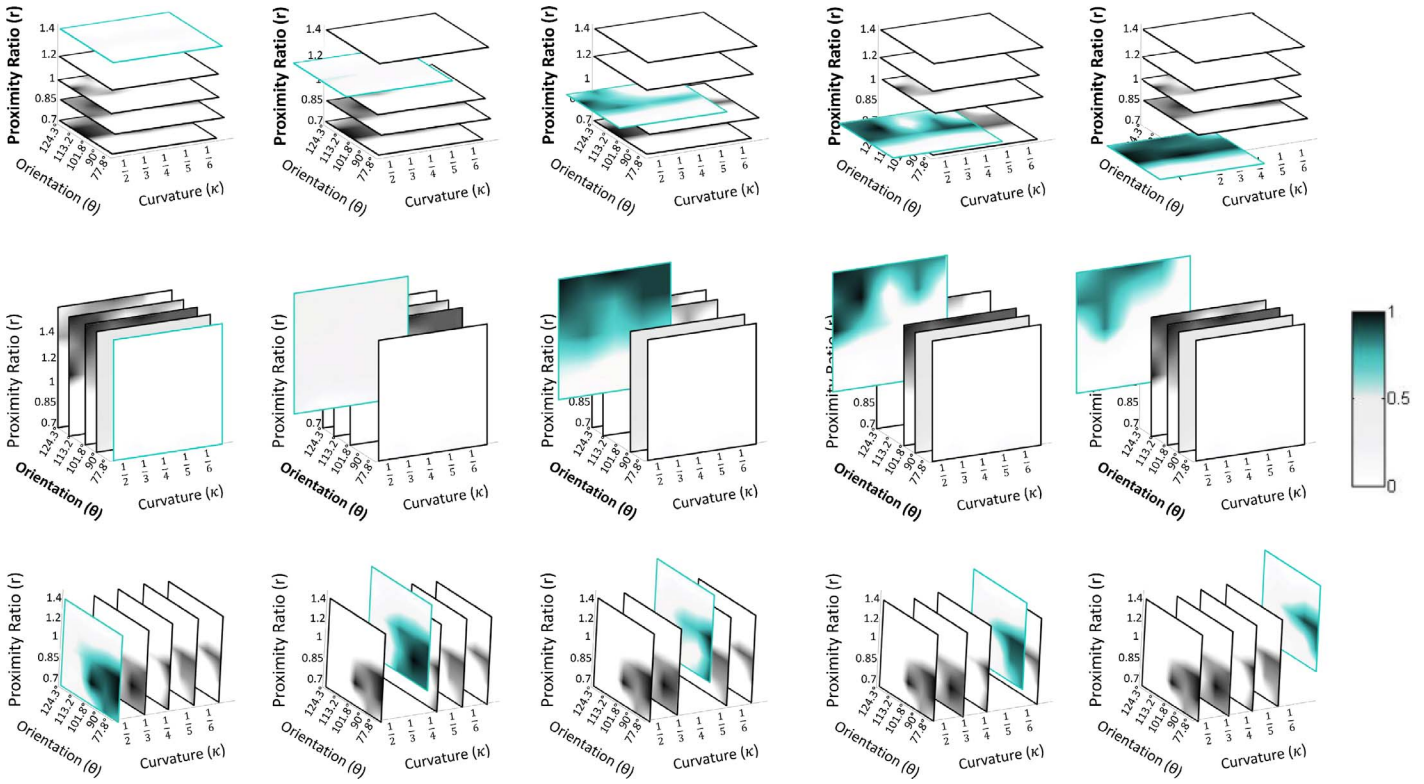

Figure 17. A graphic representation of the differences between relatability theory and the current study. Zero represents full agreement; one represents full disagreement. For clarity, only disagreement levels higher than 0.5 are show. Any nonwhite point in the resultant maps indicates a stimulus condition for which relatability and our extended model generate contradicting perceptual results.

(compared to 92.8% of our model) vis-à-vis the experimental data.

Clearly, our work is only a first step in understanding the grouping problem (and thus visual completion in general) and those stimulus factors that drive it. The presented methodology, however, is easily extensible to include additional stimulus properties, both local (e.g., the relative angle between the inducer and the contour of the occluder) and global (e.g., the shape of the occluder as a whole or the general orientation of the stimulus on the retina) as well as the effect of top-down processes (e.g., prior knowledge familiarity, expectations) on the perceptual outcome and exploration of stimuli that better represent everyday or natural perceptual tasks. Furthermore, analysis of participant reaction time may provide some insights about the perceptual importance of the different stimulus factors during the grouping and the completion process as well as its dynamics. All these are part of our upcoming work.

Keywords: grouping, inducers, curve completion, relatability theory, perceptual organization, proximity ratio, curvature, good continuation

Acknowledgments

This research was supported in part by the by the Israel Science Foundation (ISF Grant 259/12 and ISF FIRST/BIKURA Grant 281/15) and the European Commission (Horizon 2020 grant SWEEPER GA no 644313). We also thank the Frankel Fund and the Helmsley Charitable Trust through the ABC Robotics Initiative, both at Ben-Gurion University of the Negev.

Commercial relationships: none.
Corresponding author: Ben-Shahar Ohad.
Email: ben-shahar@cs.bgu.ac.il.
Address: Department of Computer Science and Zlotowski Center for Neuroscience, Ben Gurion University of the Negev, Be'er-Sheva, Israel.

References

Beck, D. M., & Palmer, S. E. (2002). Top-down influences on perceptual grouping. *Journal of Experimental Psychology: Human Perception and Performance, 28*(5), 1071–1084.

Ben-Shahar, O., & Ben-Yosef, G. (2015). Tangent bundle elastica and computer vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 37*(1), 161–174.

Ben-Shahar, O., & Zucker, S. (2004). Sensitivity to curvatures in orientation-based texture segmentation. *Vision Research, 44*(3), 257–277.

Ben-Yosef, G., & Ben-Shahar, O. (2012a). Tangent bundle curve completion with locally connected parallel networks. *Neural Computation, 24*(12), 3277–3316.

Ben-Yosef, G., & Ben-Shahar, O. (2012b). A tangent bundle theory for visual curve completion. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 34*(7), 1263–1280.

Bhatt, R., Carpenter, G. A., & Grossberg, S. (2007). Texture segregation by visual cortex: Perceptual grouping, attention, and learning. *Vision Research, 47*(25), 3173–3211.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94*(2), 115–147.

Coren, S. (1972). Subjective contours and apparent depth. *Psychological Review, 79*(4), 359–367.

DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How does the brain solve visual object recognition? *Neuron, 73*(3), 415–434.

do Carmo, M. (1976). *Differential geometry of curves and surfaces*. Upper Saddle River, NJ: Prentice-Hall, Inc.

Elder, J., & Goldberg, R. (2002). Ecological statistics of Gestalt laws for the perceptual organization of contours. *Journal of Vision, 2*(4):5, 324–353, doi:10.1167/2.4.5. [PubMed] [Article]

Fantoni, C., Bertamini, M., & Gerbino, W. (2005). Contour curvature polarity and surface interpolation. *Vision Research, 45,* 1047–1062.

Feldman, J., Singh, M., Briscoe, E., Froyen, V., Kim, S., & Wilder, J. (2013). An integrated Bayesian approach to shape representation and perceptual organization. In S. Dickinson & Z. Pizlo (Eds.), *Shape perception in human and computer vision* (pp. 55–70). London: Springer.

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local association field. *Vision Research, 33*(2), 173–193.

Froyen, V., Feldman, J., & Singh, M. (2010). A Bayesian framework for figure-ground interpretation. In J. D. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel, & A. Culotta (Eds.), *Advances in neural information processing systems* (pp. 631–639). La Jolla, CA: NIPS.

Fulvio, J., Singh, M., & Maloney, L. (2008). Precision and consistency of contour interpolation. *Vision Research, 48,* 831–849.

Geiger, D., Kumaran, K., & Parida, L. (1996). Visual organization for figure/ground separation. In *Computer vision and pattern recognition, 1996. Proceedings CVPR'96, 1996 IEEE Computer Society Conference on* (pp. 155–160). New York: IEEE.

Guttman, S., & Kellman, P. (2004). Contour interpolation revealed by a dot localization paradigm. *Vision Research, 44,* 1799–1815.

Hladky, R., & Pauls, S. (2010). Minimal surfaces in the roto-translational group with applications to a neuro-biological image completion model. *Journal of Mathematical Imaging and Vision, 36,* 1–27.

Jacobs, D. W. (1988). The use of grouping in visual object recognition. Technical Report 1023, MIT Artificial Intelligence Library. Available at https://dspace.mit.edu/bitstream/handle/1721.1/6967/AITR-1023.pdf?sequence=2

Kellman, P., & Shipley, T. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology, 23,* 141–221.

Lugo, J. E., Schmiedeler, J. P., Batill, S. M., & Carlson, L. (2015). Quantification of classical Gestalt principles in two-dimensional product representations. *Journal of Mechanical Design, 137*(9), 094502.

Maertens, M., & Shapley, R. (2008). Local determinants of contour interpolation. *Journal of Vision, 8*(7): 3, 1–11, doi:10.1167/8.7.3. [PubMed] [Article]

Ming, Y., Li, H., & He, X. (2016). Contour completion without region segmentation. *Transactions on Image Processing, 25*(8), 3597–3611.

Nayar, K., Franchak, J., Adolph, K., & Kiorpes, L. (2015). From local to global processing: The development of illusory contour perception. *Journal of Experimental Child Psychology, 131*, 38–55.

Palmer, S. (1999). *Vision science: Photons to phenomenology*. Cambridge, MA: The MIT Press.

Parent, P., & Zucker, S. W. (1989). Trace inference, curvature consistency, and curve detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 11*(8), 823–839.

Pomerantz, J. R., Goldberg, D. M., Golder, P. S., & Tetewsky, S. (1981). Subjective contours can facilitate performance in a reaction-time task. *Perception & Psychophysics, 29*(6), 605–611.

Pomerantz, J. R., & Kubovy, M. (1986). Theoretical approaches to perceptual organization: Simplicity

and likelihood principles. *Organization, 36*(3), 36–1 –36-46.

Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory, 2*(5), 509–522.

Prazdny, K. (1985). On the nature of inducing forms generating perceptions of illusory contours. *Perception & Psychophysics, 37*(3), 237–242.

Reisberg, D. (2013). *The Oxford handbook of cognitive psychology*. Oxford, UK: Oxford University Press.

Shipley, T. F., & Kellman, P. J. (1992). Perception of partly occluded objects and illusory figures: Evidence for an identity hypothesis. *Journal of Experimental Psychology: Human Perception and Performance, 18*(1), 106–120.

Singh, M. (2004). Modal and amodal completion generate different shapes. *Psychological Science, 15*(7), 454–459.

Spelke, E. S. (1990). Principles of object perception. *Cognitive Science, 14*(1), 29–56.

Summerfield, C., & De Lange, F. P. (2014). Expectation in perceptual decision making: Neural and computational mechanisms. *Nature Reviews Neuroscience, 15*(11), 745–756.

Takeichi, H. (1995). The effect of curvature on visual interpolation. *Perception, 24,* 1011–1020.

Takeichi, H., Nakazawa, H., Murakami, I., & Shimojo, S. (1995). The theory of the curvature-constraint line for amodal completion. *Perception, 24,* 373–389.

Unuma, H., Hasegawa, H., & Kellman, P. J. (2010). Spatiotemporal integration and contour interpolation revealed by a dot localization task with serial presentation paradigm. *Japanese Psychological Research, 52*(4), 268–280.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychological Bulletin, 138*(6), 1172–1217.

Ward, E. J., & Chun, M. M. (2016). Neural discriminability of object features predicts perceptual organization. *Psychological Science, 27*(1), 3–11.

Wertheimer, M. (1923). *Laws of organization in perceptual forms. A source book of Gestalt psychology*. London: Routledge & Kegan Paul.

Witkin, A., & Tenenbaum, J. (1983). On the role of structure in vision. In J. Beck, B. Hope, & A. Rosenfeld (Eds.), *Human machine vision* (pp. 481–542). Cambridge, MA: Academic Press.

Appendix: Real-valued model

The model obtained in Equation 4 is based on the index or "level" of the parameters (as defined in Figure 3). This was necessary in order to make the parameters commensurable despite having different units and overwhelmingly different ranges and scale of values. However, it is easy to transform this representation to one in which the true/real values of the parameters can be plugged in for prediction of the outcome in novel conditions. More specifically, the conversion from its level value (one through five) to the real-valued, unitless proximity ratio is

$$R = 0.174 \cdot r + 0.592. \quad (5)$$

The mapping that converts the relative orientation level (1 through 5) to real-values relative orientation in *degrees* is

$$\Theta = 11.6 \cdot \theta + 66.5, \quad (6)$$

while the mapping of curvature level (one through five) to real-value curvature in $length^{-1}$ units is simple:

$$\Gamma = \frac{1}{\kappa + 1} \quad (7)$$

Plugging these expressions into the model from Equation 4 produces the final real-values model:

$$\begin{aligned} P[\text{perceive two shapes}] = {} & 1.115 + 0.804 \cdot R \\ & - 0.0120 \cdot \Theta \\ & - 0.033 \cdot \frac{1}{\Gamma} \end{aligned} \quad (8)$$