Postscript: Identity and Constraints in Models of Object Formation

Philip J. Kellman
University of California, Los Angeles

Patrick Garrigan
University of Pennsylvania

Thomas F. Shipley
Temple University

Brian P. Keane
University of California, Los Angeles

As this exchange concludes, we believe that the account of interpolation and object formation proposed by Kellman and Shipley (1991), further developed in recent years (Kellman, 2003; Kellman, Gutman, & Wickens, 2001), and most recently extended to 3-D interpolation (Kellman, Garrigan, & Shipley, 2005) and spatiotemporal object formation (Palmer, Kellman, & Shipley, 2006), remains viable. Here we briefly note some progress in this discussion, including positions taken by Anderson (2007a) that have since been abandoned. We address the new positions that Anderson (2007b) takes, which now focus on interpolations that switch between modal and amodal appearance, data on interpolated contour shape, evidence and methodological concerns about early interpolation, and physiological evidence.

Anderson’s initial commentary attacked our theory of relatability, yet the arguments against these empirically supported geometric constraints seem to have been dropped, and no other account has been offered of which geometric relations of edges in two or three dimensions supports interpolation in unit formation. Arguments about the identity hypothesis have changed significantly. For example, at first, Anderson (2007a) cited results suggesting that luminance constraints can block modal but not amodal inter-
polation (Anderson, Singh, & Fleming, 2002). We pointed out a confound in the experimental design. Anderson (2007b) now states that this confound was controlled “as closely as possible,” which from our point of view is consistent with the small effect observed. More simply, Anderson’s adjusted position—that luminance constraints do not block modal completion but merely weaken its appearance—is consistent with the version of identity we have endorsed since 2001 (Kellman, 2003; Kellman, Garrigan, & Shipley, 2005; Kellman et al., 2001). The strongest arguments for identity remain the logical implications of certain observed phenomena. To challenge these (including Petter’s effect, quasimodal contours, and crossing interpolations), Anderson (2007a) made ad hoc claims about various decays of interpolation strength. We pointed out problems and inconsistencies, and now these issues have been dropped. The only consistent way we know to explain crossing interpolations, depth spreading, and quasimodal contours is that modal and amodal completion share a common underlying interpolation process; the terms modal and amodal designate appearances in the final percept, not separate interpolation processes. In a new criticism, Anderson (2007b) disparages our model as “feedforward.” How this description fits a model that proposes an early, overgenerating interpolation process followed by confirming, weakening, or deleting of some interpolations on the basis of other scene constraints is unclear. In fact, it appears that Anderson gets the argument reversed: 

If the relative depths of contour segments are the very ingredients of the contour interpolation process, then why isn’t there sufficient information to determine the modal and amodal status of the interpolated contours from the outset? (Anderson, 2007a, p. 480)

If all scene information is present when interpolation occurs, what is the role of feedback? Anderson describes our model as inefficient but offers in support neither a formal rationale nor even a shareable intuition. Given the new issues, different emphases, and apparent inconsistencies in the latest critique, writing a comprehensive response in a short space is difficult. Anderson seems to address a number of issues by changing the subject to talk about transmittance, transparency, TAP, and CDAP. He now claims that the latter acronyms explain all sorts of things, despite being unworthy of a single mention in his first response to our manuscript. We therefore provide brief rejoinders on several matters and slightly more extended responses on issues that seem substantive and useful at this point.

Constraints on Perceived Interpolation

Anderson (2007b) acknowledges that clear evidence of interpolated contours that are modal through part of their length and amodal in other parts would provide support for an identical contour interpolation process. We published such evidence in 1998 (Kellman, Yin, & Shipley, 1998), and furnished clear examples in other work (e.g., Kellman, Garrigan, & Shipley, 2005; Kellman, Garrigan, Shipley, & Keane, 2007). Rubin (2001) also published data on such interpolated contours (Figure 14C, p. 356). These published examples should be sufficient, but Anderson has now raised the bar. He says that phenomenal impressions of interpolation are absent when two occluding dots are placed at the endpoints of amodal contours. This new example is interesting, as there appears to be a constraint on how many intersecting contours an interpolated edge can cross. Whereas in our quasimodal displays and Rubin’s 2001 example, quasimodal contours appear to survive a single occluder, Anderson’s (and a different display studied by Rubin; see his Figure 14A) have two. However, Figure P1 shows that Anderson is incorrect in stating that only an amodal contour is prevented from switching to modal. The source of the effect is that there are two occluding figures, rather than one. It is important to note that this constrains contours that begin either as modal or amodal completion, as shown in Figure P1B. This constraint regarding two occluders (or number of crossing edges) affects both modal and amodal contours and is therefore compatible with the identity hypothesis. We can go a step further. Figure P1C shows that an inducer placed in between the two occluders can re-establish interpolation. This indicates that interaction across one occluder is possible, as has been previously shown with quasimodal contours and by Rubin (2001). Another set of issues regarding constraints on interpolation involves crystalline interpolations. We offered these to show that modal completion can occur despite the absence of the luminance relations that Anderson (2007a) previously claimed to be crucial. Anderson (2007b) changes the subject to write that his work on transparency explains why these displays form and why they look transparent. He appears to believe that being previously aware of these displays somehow negates the fact that they contradict his claims. If transparent interpolated objects form in the absence of the luminance relations that Anderson has deemed necessary, just what is the luminance constraint on modal completion? Why and when interpolation may involve a transparent or detached surface was first discussed by Shipley and Kellman (1994). The connection of visible fragments across gaps in these displays is predicted by our model. What Anderson offers as an explanation is beside the point (at least for contour interpolation). Consider our Figure 3B (Kellman et al., 2007). It shows that disrupting relatability by shifting the inducing contours eliminates unit formation (connections across gaps). These shifts break up interpolation but leave undisturbed the depth and luminance relations that are the subject of Anderson’s CDAP and TAP.

Evidence for Identity and Early Interpolation

Anderson (2007b) suggests discarding all results obtained from a frequently used paradigm that has been shown to be sensitive to contour interpolation: the fat–thin paradigm of Ringach and Shapley (1996). His reason is that in Guttman and Kellman (2005) “no difference in performance was observed for any of the displays containing relatable contours, despite clear differences in the perceived strength of completion” (Anderson, 2007b, p. 522). In other words, Anderson asserts that a method must be flawed if it produces results contradicting phenomenology. If we accept this reasoning, our hypothesis about early interpolation, a process occurring before final scene percepts are determined, would be experimentally untestable. Subscribing to this view renders every unconscious process untestable. As we do not subscribe, we tested our promiscuous interpolation hypothesis. The data clearly support what Anderson deems impossible: Early interpolation occurs for edges where tangent discontinuities and relatability are present, even though some of these edges (e.g., in outline displays) do not produce perception of interpolated contours. Anderson neglects to mention, but helpfully reprints, the data from some of the various
control groups that make this case. The evidence shows that sensitivity \(d')\) enhancements occur, relative to the control group, for all displays with tangent discontinuities and relatable contours. (Curiously, Anderson asserts that our L-shaped inducers lack tangent discontinuities, but this is incorrect.) The full set of studies controls for a number of alternative explanations, including all the ones Anderson mentions. The performance advantages for relatable contours are not due to grouping effects, and they survive when cross-trial influences (priming) caused by mixing stimuli are removed. On this topic, Anderson reprints (his Figure 10) Guttman’s data from an experiment (Guttman & Kellman, 2005) comparing subjects who saw only outline displays with those who saw both outline and filled displays. This comparison allowed Guttman to clearly identify small cross-stimulus priming effects (remarkably consistent across all five inducer rotations at about 0.5 \(d'\) units). Inexplicably, Anderson asserts that this result “shows that such strategies can completely obscure differences in perceived completion strength” (Anderson, 2007b, p. 522). He must have failed to observe that in all cases, the relatable outline stimuli outperformed their nonrelatable controls in robust fashion. This might have been clearer in Anderson (2007b) if he had not chosen to reproduce only two thirds of a figure from the paper. The missing panel (shown here in Figure P2) shows the data from subjects who viewed only outline displays and compares them with data from subjects who viewed only filled displays. The interpolation effects are obvious in both; in this case, where cross-trial influences between filled and outline displays are precluded, outline displays show clear interpolation effects but reliably less than do filled displays. A separate assertion by Anderson is that differences in edge position between conditions could account for our results, despite control conditions showing that this is not the case. Guttman and Kellman (2005) also carried out a crucial test validating that the fat-thin paradigm actually measures interpolation effects. One way to do this is to show that two manipulations that independently eliminate contour interpolation have no additive effects on performance (because of factors unrelated to interpolation) when both are included. This prediction was confirmed with rotating (Manipulation 1) and laterally shifting (Manipulation 2) inducers (Experiment 3). In sum, each and every result in Guttman and Kellman’s studies supports a theory of early interpolated connections that form and can have processing effects but that do not all survive into final scene representations. This notion, by the way, is not unique to our model: In interpolation models, it was explicitly implemented in Grossberg and Mingolla (1985) and subsequent work by those investigators, and a similar idea was invoked by Marr (1982) to explain other perceptual phenomena.

Shape Issues in Interpolation

Do modal and amodal completion generate different shapes? Earlier, we and Albert (2007) independently analyzed Anderson’s serrated edge, star, and cross displays, and we both concluded that the interesting effects do not arise from separate completion processes. Anderson (2007a) claimed, for example, that differences in modal and amodal completion underlie the effect demonstrated in his cross displays. Albert showed that these effects occur with transparency displays in which there is no completion at all! (Incidentally, the effect also shows up in interpolation displays when both of the crossing interpolations are amodal or both are modal.) So the effect cannot derive from a difference in processes of modal and amodal completion. Anderson’s (2007b) backup position is a new theory that opacity and transparency are differences not in kind, but in degree. Such a claim may be reasonable, but it comes at a high price: This claim is incompatible with the idea of distinct modal and amodal completion processes (they would be matters of degree of a unified process). So, in essence, Anderson does explain Albert’s effect, but ironically by accepting

---

**Figure P1.** Displays showing effects of multiple interruptions along an interpolation path. All displays are stereopairs (free-fuse by crossing the eyes). (A) Ordinary illusory figure, with clear curving top and bottom edges. (B) Placement of two dots (projectively) in the interpolation path interrupts perceived contour connection for top and bottom edges. Anderson (2007b) claimed that such a blocking effect is specific to contours that begin as amodal. This demonstration shows that it occurs with interpolations that begin as modal. There is a constraint on perceived interpolation but not an amodal or modal difference. (C) Addition of an inducer in the empty space re-establishes robust perceived interpolation. This display confirms other observations and data (e.g., Rubin, 2001) indicating that interpolated contours can survive one intervening occluder (as in quasimodal completion). (D) This display shows that the added elements in Panel C do not by themselves produce the robust curved interpolations seen in that display.
the identity hypothesis. In his earlier commentary, Anderson described as most compelling Singh’s (2004) work claiming that modal and amodal interpolation produce different contour shapes. We noted that Singh used a subjective method (adjustment) and that at least two processes were likely at work in his amodal conditions. One is a local contour interpolation process; the other is recognition from partial information (RPI). As we have described earlier (Kellman, 2001; Kellman, Garrigan, & Shipley, 2005) RPI is a more cognitive tendency to report familiar or symmetric forms behind an occluder. In all of Singh’s displays, straight edges could meet at a corner behind an occluder forming a familiar or symmetric polygon. We suspected this effect, given the direction of Singh’s results: He stated that amodal completions tend to be more “cornerlike.” We recently were able to look further into the data of Singh (2004). Singh presented illusory and occluded displays in a within-subjects design and had subjects make settings of how much smoothing (deviation from a sharp corner) they perceived in interpolated contours. Displays were tested in two experiments with four gap sizes and two contour angles. Each display was presented five times per subject. From our investigation of the data, we would argue strongly against drawing conclusions about shapes of interpolated contours from these experiments. The reasons have to do with our earlier hypothesis about RPI effects, as well as with variability, within and between subjects. First, in the two experiments, 7 subjects (3 in Experiment 1 and 4 in Experiment 2) reported that the occluded contours met at a sharp corner for more than 72% of the trials for all displays in the amodal condition. Only one subject (in Experiment 2) gave such a high percentage of corner responses for modal displays. Responding that the shape forms a corner is inconsistent with other measurements and most models of contour interpolation (e.g., Fantoni & Gerbino, 2003; Kellman & Shipley, 1991) but is consistent with RPI. The substantial subgroup of subjects giving mostly “corner” responses in the amodal condition, as predicted by RPI, suggests that this response tendency competes in this paradigm with local interpolation tendencies. It is not a priori obvious how subjects will merge these influences, but in the event that some simply respond “corner” to most amodal displays, as occurs in this data set, it supplies clear evidence of the RPI tendency. Subjects who make mostly “corner” responses should not be simply averaged in with subjects who make responses indicating smooth interpolation, but that is how those data were treated. Not surprisingly, the means of the subgroup of corner responders differed substantially from those of the remaining subjects, with the differences for most data points being larger than those reported in the group data between amodal and modal conditions. The data further suggest that subjects who were not mostly corner responders were probably balancing between interpolation (smooth completion) and RPI (corner completion) for amodal displays. In Experiment 1, corner responses occurred 201 times for amodal displays and 73 times for modal ones. In Experiment 2, corner responses occurred 231 times for amodal and 120 times for modal ones. Again, as predicted, RPI acted predominantly in amodal cases.

We thank Manish Singh for graciously sharing his data. Singh stands by his interpretation of the data for the reasons outlined in Anderson’s (2007b) postscript. We also note that our analysis does not deny that the two groups in Singh’s (2004) studies produced reliably different patterns of data; our concern is whether the measurements in these studies reveal the shapes of amodal and modal contours and whether they show a meaningful difference between modal and amodal shapes.
Competing interpolation and RPI response tendencies are supported by the data, but there is another concern regarding whether these experiments converged on specific shapes for modal or amodal contours. Looking at the data, we noted that within the five settings by a subject for each display, there was high variability. On a number of occasions, subjects gave the exact same stimulus the minimum smoothing response (most cornerlike) on one trial and the maximum (most rounded) on another. This within-subject variability (variability in the set of trials that gives a single data point for a single stimulus for one subject) is not reflected in the error bars in the published data; these show variability across subjects. Figure P3 conveys some of what is going on in individual data. Figure P3A displays histograms of all individual responses that contributed to single published data points for one modal and the comparable amodal stimulus, specifically the stimuli with smallest gap (Figure P3A, left) and the largest gap (Figure P3A, right) in Experiment 2. Responses to the modal version of each display are shown above responses to the amodal version. The three thin vertical lines in each panel show the mean (middle line) and standard error (flanking lines) reported by Singh in the published graphs. (B) Individual subject responses for all separation values from the 75° turn-angle condition in Experiment 2.

Figure P3. Variability in individual responses for stimuli in Singh (2004). Each display shows a histogram plotting the frequency of shape responses (given by subjects as adjustment of interpolated contour shape of a comparison stimulus and encoded on a normalized scale of contour smoothing). Subjects made five responses to each of two turn angles tested at the separation value shown; these responses were averaged together to get each subject’s score in the group data. Here we show each of the five individual responses for all subjects as separate points. (A) All responses that contributed to the published data point for the smallest gap (left) and the largest gap (right) in Experiment 2. The blue line marks the score shown in the published data (M) and the red lines show the error bars (± 1 SE). It can be seen that the spread of actual responses offers little

Each histogram includes all trial responses that contributed to a single published data point in Singh (2004). All data points in that article were combined across two turn angles tested. Although we follow this grouping of the data in Figure P3A, our observations also hold when data for each turn angle are separated (Figure P3B).
support for the idea that the mean value reflects the perceived shape of an interpolated contour. In fact, for the stimulus in the left half of Figure P3A, 94.3% of all subjects’ settings for amodal displays fell outside the standard error bars. For modal displays, this value was 87.1%. The large variability is not unique to this particular case. For the stimulus in the right half of Figure P3B, 91.4% of all responses fell outside of the error bars for amodal displays, and 87.9% for modal displays. Although there is no well-defined expectation for how many responses should fall within the error bars that describe intersubject variability, the point is that the means were generated from a highly varying and inconsistent set of underlying individual responses. Individual subject responses for all separation values from the 75° turn-angle condition in Experiment 2 are shown in Figure P3B. In recent work (e.g., Fulvio, Singh, & Maloney 2006), Singh and colleagues have used consistency of responding as an indicator of the conditions under which interpolation really occurs. Indeed, data suggest that when tested with methods of suitable power, human contour interpolation is consistent and precise, within and across subjects (e.g., Guttmann & Kellman, 2004). Far from providing, as Anderson (2007b) puts it, a compelling reason to believe in separate modal or amodal completion processes (or even, as Singh suggested, a version of the identity hypothesis with a shape parameter), the data of Singh (2004) simply do not converge on consistent results relating to perceived shapes of interpolated contours.

Physiology and Identity

Perhaps it signifies progress in this exchange that the arguments have changed. Anderson’s initial commentary (Anderson, 2007a) contained a remarkable number of separate arguments but not a single reference to physiological data. With our response in hand, Anderson’s (2007b) latest position is that “Kellman and colleagues’ model . . . must ultimately be resolved by physiological data” (p. 512). We believe that such data are important, but they are by no means the only source of information about visual processing. This general point is specifically true in research on interpolation in object formation, as physiological data have not yet converged on a clear understanding of the loci of various functions, much less on the details of relevant computations. Yet the data already offer important clues. We do not understand Anderson’s (2007b) pronouncement that “all of the extant evidence seems very much against the identity hypothesis” (p. 512). Murray, Foxe, Javitt, and Foxe (2004) explicitly interpreted their data in the lateral occipital cortex (LOC) and parietal regions as strongly supporting the identity hypothesis. Anderson’s characterization also incorrectly summarizes evidence in V1 and V2. For example, Lee and Nguyen (2001) found individual neurons that responded to both modal and amodal interpolations reliably more than they responded to a control stimulus. Although amodal responding was on average lower than modal, amodal displays produced the same spatial precision as real and illusory contours. Responses to both were also clearly evident in averages over many cells, especially in superficial layers of V2 over the first 200 ms. These results are on their face clearly consistent with our theory. Anderson argues that “it is highly unlikely that the identical responses observed by Murray et al. (2004) have anything to do with contour interpolation processes” (p. 512). The basis of his argument is functional magnetic resonance imaging work by Stanley and Rubin (2003) on LOC, in which similar activation of LOC was found in response to illusory contour stimuli and variants that disrupted clear contour perception by rounding tangent discontinuities. Stanley and Rubin (2003) called these latter displays salient region stimuli, because the rounded inducers still seemed to partly define a region. Consistent with some proposals in computer vision, they argued that perhaps LOC determines salient regions and sends this information to V1 or V2, where contour interpolation may take place. This interesting proposal faces several difficulties. First, Mendola, Dafe, Fischl, Liu, and Tootell (1999) found illusory contour responses in LOC not only with Kanizsa-style inducers, but also with displaced grating stimuli. The latter do not encompass salient regions. More important, the hypothesis of Stanley and Rubin is a temporal one, suggesting that vaguely defined regions precede contours. Their functional magnetic resonance imaging methods lacked the temporal resolution to determine whether such a sequence actually occurs in LOC. Electrophysiological methods, such as those of Murray et al. that employ high-density visual evoked potential recordings, topographic analyses, and local autoregressive average source analysis allow better temporal resolution. More recent results using these techniques (Shpaner, Murray, & Foxe, 2007) clarify the picture considerably. They used the illusory contour displays and SR displays of Stanley and Rubin (2003) and showed much stronger initial responses to illusory contour displays than to SR displays as well as a difference in response topography thereafter. Investigation of these issues will no doubt continue. The immediate point is that the available evidence does not support Anderson’s dismissing of LOC as a site of contour interpolation. So we are back to the point we made earlier: The results of Murray et al. (2004) constitute strong evidence for the identity hypothesis in humans, with methods superior in many respects to those in other studies. Finally, we note that Murray et al. (2004) also found early and comparable illusory and amodal responding, compared with a control group, in posterior parietal areas. Stanley and Rubin (2003) did not study these areas. These latter results are consistent with an argument advanced by Kellman, Garrigan, Shipley, Yin, & Machado (2005). They argued from their data that 3-D interpolation would require mechanisms that possess information about depth intervals and contour slant in the world. Specifically, they noted as a candidate area the caudal intraparietal sulcus, because both single-unit recording in monkeys (Sakata, Taïra, Kusunoki, Murata, & Tanaka, 1997) and functional magnetic resonance imaging in humans (Shikata et al., 2001) suggest that this region may represent real slant specified by binocular or other depth cues.

Processes of perceptual organization, and object perception in particular, comprise central topics in understanding the mind, and they pose no shortage of complex challenges to researchers. Our consideration of issues in this dialogue gives us confidence that the kind of model we have proposed has proven useful in several respects. It connects diverse phenomena that share common underlying mechanisms, explains the geometric relations that give rise to interpolation, and defines an architecture that clarifies relationships of local and global, as well as contour and surface, influences in object formation. We assume continuing psychological, computational, and physiological research will improve these ideas but also will incorporate key features that evidence already indicates should be included in accounts of object perception.
References