Hierarchical Grouping of Simple Visual Scenes

Patrick J. Rice[1], Jiayue (Linda) Mao, Ziying (Ariana) Zhu, Emily Wu, Michael D. Byrne[2]
[1]pjr@rice.edu
[2]byrne@rice.edu
Rice University, Department of Psychology, 6100 Main Street
Houston, TX 77005 USA

Abstract

Human visual grouping processes consolidate independent visual objects into grouped visual features on the basis of shared characteristics; these visual features can themselves be grouped, resulting in a hierarchical representation of visual grouping information. This "grouping hierarchy" promotes efficient attention in the support of goal-directed behavior, but improper grouping of elements of a visual scene can also result in critical behavioral errors. Understanding of how visual object/features characteristics such as size and form influences perception of hierarchical visual groups can further theory of human visual grouping behavior and contribute to effective interface design. In the present study, participants provided free-response groupings of a set of stimuli that contained consistent structural relationships between a limited set of visual features. These grouping patterns were evaluated for relationships between specific characteristics of the constituent visual features and the distribution of features across levels of the indicated grouping hierarchy. We observed that while the relative size of the visual features differentiated groupings across levels of the grouping hierarchy, the form of visual objects and features was more likely to distinguish separate groups within a particular level of hierarchy. These consistent relationships between visual feature characteristics and placement within a grouping hierarchy can be leveraged to advance computational theories of human visual grouping behavior, which can in turn be applied to effective design for interfaces such as voter ballots.

Keywords: visual grouping; human factors; ballot design; UI

Introduction

To support effective processing of behaviorally-relevant portions of visual scenes, the visual system must group sets of visual objects into features where object membership is based on some meaningful characteristic. For example, while the three lights of a traffic light are of different colors, they are vertically or horizontally aligned and contained within a rectangular bounding box, and are perceived of by humans to belong to a functional unit that demands attention while engaged in a complex behavioral task (that is, driving). Humans also easily distinguish between two traffic lights at the same intersection on the basis of orientation. This grouping of independent objects into features (three aligned lights grouped into a traffic light), and of features into larger, more inclusive features (multiple traffic lights facing with the same orientation grouped for a particular direction of travel, and traffic lights facing different directions, but within some proximity of one another, grouped into an intersection) implies that human perception of visual groups is hierarchical in nature.

Understanding how various stimulus characteristics and relationships inform human visual grouping perception is essential for effective communication of designer intent to users in usability scenarios. For example, if a specific visual characteristic is consistently associated with grouping of visual elements that possess that visual characteristic at a particular level of a hierarchy, that tendency can be leveraged to guide user's attention effectively in visual scenes that contain unfamiliar visual elements and features. Another characteristic may consistently be perceived of as distinguishing separate features that reside at a specific level of a grouping hierarchy, allowing users to distinguish otherwise similar visual features.

One example where understanding of how specific visual characteristics contribute to hierarchical group perception is in the design of voting ballots. In recent years, there have been a number of examples of poor ballot design leading to relatively high likelihood of voter error, such as omitting a vote for a particular contest, or voting for two candidates in the same contest (Norden, Kimball, Quesenbery, & Chen, 2008). In one case, it is very likely that the outcome of a congressional election was changed due to undervoting on a specific contest; this undervoting was related to the arrangement of contests and instructional elements on the first page of the ballot, which immediately preceeded the page of the undervoted contest (Greene, 2010). The arrangement of the instructional and contest-related elements on the initial page informed voter's expectations regarding the position of contests on subsequent pages; the position of the undervoted race violated voter's expectations, appearing in a location that formerly contained instructional material.

Most elections in the United States are administered at the county level, and there are over 3,000 counties in the U.S. Within each county, there are often hundreds of different precincts, each with a slightly different ballot style, meaning that for each national election, tens–perhaps hundreds–of thousands of different ballot designs are deployed. County clerks simply do not have the expertise or resources to conduct a detailed usability test of every ballot they present to voters on election day. Computational cognitive models of human visual grouping processes have the potential to address this shortcoming by providing a widely accessible usability testing tool that can be applied to a large number of ballot designs in a fast and consistent manner.

The ACT-R cognitive architecture has successfully been used in multiple human factors applications (Salvucci, 2006; Byrne & Kirlik, 2005), and can be leveraged to design a usability testing tool which utilizes knowledge on human visual grouping perception to provide predictions of voter behavior across a wide range of ballot designs. One example is a visual grouping algorithm for the ACT-R cognitive architecture which uses the proximity of visual objects in a scene to provide a prediction of group membership across elements of the visual scene (Lindstedt & Byrne, 2018). While the existing model makes qualitatively sensible predictions of group membership, it is incapable of representing the sort of nested group features that hierarchical grouping requires. For example, the existing model identifies the names of the candidates of a particular contest as belonging together in a group, and the buttons allowing for selection of each of those candidates as belonging to another group, but does not group the name of a candidate to the corresponding button for their selection, a critical understanding that guides voter behavior.

While proximity is known to be a major contributor to group creation (Wagemans et al., 2012), it is likely that other characteristics contribute to hierarchical visual grouping. To more fully characterize contributors to the hierarchical nature of human visual grouping, the present study aims to determine characteristics of objects and features of simple visual scenes that are associated with specific patterns of hierarchical grouping. Visual object/feature characteristics of size, alignment, form, and object/feature status will be related to the relative placement of visual objects and features into a hierarchical set of groups. To do so, participants were shown stimuli containing a limited set of stimulus objects and features that had consistent relationships in size, alignment, and form, and allowed to make free-response indications of their perception of grouping of the scene. These free response grouping patterns were coded according to a labeling system identifying the grouping together of specific stimulus elements, and these labels were then evaluated for relationships to the characteristics of the objects and features that drove the grouping.

Method

Participants

For this study, 39 individuals (33 female, mean age = 19.5, *SD* = 0.97) were recruited from the Psychology Research Pool at Rice University. All participants received credit towards a course requirement for completing the study. The experimental protocol was approved by Rice University's Institutional Review Board.

Design

To evaluate visual grouping behavior, we developed a free-response task paradigm which allowed participants to indicate subjectively perceived hierarchical groups of simple visual objects and features. Participants were shown a set of 34 stimuli containing a number of simple visual features drawn from a set of three unique features (see Figure 1 for examples of these three features) and asked to provide free-response groupings of the visual objects/features contained in the stimuli.

The study design was a one-way between-subjects design with two conditions to which participants were randomly assigned in an alternating manner: 1. top-down and 2. bottom-up instruction formats (see next paragraph for details). A set of eight colored pens were provided to participants alongside a packet in which each of the 34 stimuli were printed on separate pages; the order of stimuli within each packet was randomized. Participants were instructed that their task was to indicate their perception of the grouping of visual objects in each stimulus by circling each perceived group in pen. The task instructions indicated to participants that a single visual object can belong to a group by itself, and that groups of objects may be contained by larger groups (that is, groups of visual objects can be grouped together into "meta-groups"). The instructions also indicated that multiple independent groups could exist on the same level of hierarchy, and to assume that all objects within a given stimulus belong to one large "meta-group" that did not need to be marked.

All participants were instructed to use different colors to circle groups on different levels of the perceived hierarchy (e.g., all groups on the one level of the hierarchy were marked with one color, all groups on another level of the hierarchy were marked with a different color, and so forth); however, the two conditions of the experiment were defined by how participants were instructed to indicate groups that were perceived to exist on different levels of a hierarchy. One group of participants (n = 19) received instructions which emphasized first circling all groups that were perceived to be on the "lowest" level of the hierarchy and that could not be further subdivided into smaller groups before switching colors to indicate groups on the "next" level, and so on (the "bottom-up instruction" condition). Participants in the second condition (n = 20) received instructions which emphasized first circling all groups that were perceived to be on the "highest" level of the hierarchy and that contained the most objects, before switching colors to indicate groups perceived to be contained by these "higher-level" groups. Participants were allowed to decide which specific colors to use for different levels of hierarchy.

Both sets of instructions specified that there is no expected number of groups that the participant should identify, and that it was not required that each color of pen be used for each stimulus. The instructions also contained an example of a grouped visual scene for participants to become familiar with. This example was not included in the set of stimuli, and displayed groups residing on three levels of hierarchy (the example provided to participants is shown in Figure 2). Participants were informed that this example may not present the same groupings that they would have marked.

Stimuli and Materials

Stimuli in this task were each comprised of a number of simple visual features drawn from a set of three unique features: a 20 mm by 20 mm square, a 20 mm by 100 mm rectangle, or a 3x3 grid of equally spaced 4 mm by 4 mm squares ("small squares"), which were positioned so that set of nine "small squares" (referred to hereafter as a "nine-set") occupied the same area as that of the larger 20 mm by 20 mm square (see Figure 1 for examples). It was possible for the rectangular visual feature to be oriented either vertically or horizontally in

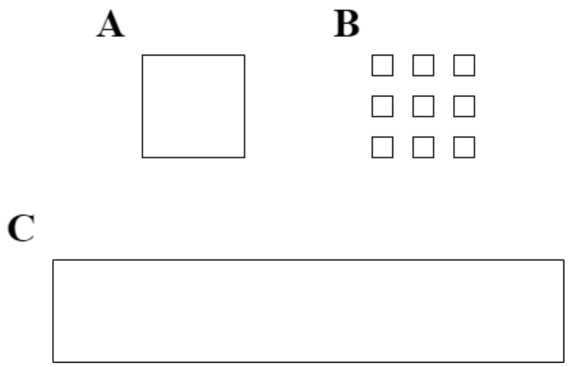

Figure 1: Each of the three visual features used to create stimuli for the task. A: the 20 mm by 20 mm square visual feature; B: the visual feature comprised of a 3x3 grid of equally spaced 4 mm by 4 mm squares (a “nine-set,” comprised of “small squares”); C: the rectangular 20 mm by 100 mm visual feature. This feature could appear oriented either horizontally or vertically in a stimulus, but not both. Image not to scale.

a given stimulus, but no stimulus contained rectangular visual features in both orientations.

Stimuli contained a minimum of four and a maximum of 81 independent visual objects (*M* = 29.6, *SD* = 22.0; note that the ”grid” visual feature is comprised of nine visual objects). When collapsing across visual features, stimuli contained a minimum of four and a maximum of nine independent visual features (*M* = 6.35, *SD* = 1.97). Six stimuli containing four visual features each were included to replicate stimuli from a previous piloting experiment in the lab; the remaining 28 stimuli were designed so that there were nine potential visual feature positions arrayed in a 3x3 grid. Squares and “nine-sets” occupied a single position each, while rectangles occupied three positions. In 24 of these stimuli, the nine potential feature positions were equidistant from one another (40mm center-to-center vertical or horizontal separation, so that the vertical or horizontal distance between the edges of two adjacent visual features was 20mm); however, in the remaining four stimuli, the feature positions were spread such that there was twice as much horizontal distance between the edges of two horizontally adjacent visual features (40 mm edge-to-edge, two stimuli) or twice as much vertical distance between the edges of two vertically adjacent visual features (40 mm edge-to-edge, two stimuli). When the distribution of visual features was not symmetrical along both horizontal and vertical axes, a 90° rotated version of the stimulus was created so that versions with both horizontal and vertical alignment were included.

The 28 total 3x3 grid stimuli were created so that there was consistent hierarchical structure between the included visual features. Specifically, we focused on three relationships: 1. The 4mm by 4mm and 20mm by 20mm squares had the same length-to-width ratio, and only differed in size; 2. the 20mm by 20mm squares and the “nine-sets“ spanned the same square area, but differed in whether they were a singleton visual object (the square) or a grouped visual feature comprised of multiple visual objects (the “nine-set”); and 3. three vertically/horizontally adjacent squares and/or “nine-sets” occupied the same area as rectangles, but were again distinguished as a visual object versus a visual feature.

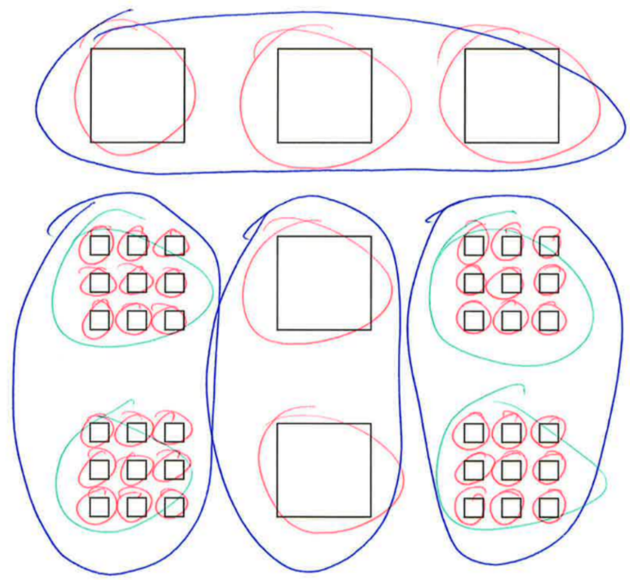

Figure 2: The example included in the set of instructions for both the “bottom-up” and “top-down” instruction conditions. Groups on three levels of hierarchy have been indicated, with red circles marking groups on the “lowest” level of the hierarchy and blue circles marking groups on the “highest” level.

Data Analysis

Data from eight participants was excluded from analysis– six participants did not complete the task, while responses from the remaining two excluded participants indicated that they did not respond in accordance with the task’s instructions. To evaluate the free response groupings of the visual objects/features for each stimulus, a coding scheme was developed. This scheme includes a set of labels to indicate the presence of particular grouping patterns (such as the circling of rows/columns, or the grouping together of a set of visual objects that form a “T” shape), as well as syntax to specify any hierarchical structure across groups that a participant’s response may have contained. Evaluation using this scheme resulted in consensus between three independent coders for all participant responses. The sets of labels categorizing each participant’s hierarchical grouping responses for each stimulus were then analyzed to identify the most frequently occurring responses for each stimulus across participants. These modal responses were further evaluated for the frequency and position in hierarchy of specific grouping patterns, and then the hierarchical positions of these patterns were compared.

Results

The major finding was that participants' free-response groupings demonstrated consistent sensitivity to two general characteristics of the structure contained within the stimuli: 1. the relative size of visual objects/features distinguished groupings across levels (smaller visual objects/features were more likely to be grouped at lower levels, while larger visual objects/features were more likely to be grouped at higher levels); and 2. the form/shape of visual objects/features was related to differentiation of groups within a level.

Across stimuli and participants, a total of 365 unique grouping patterns were observed. On average, participants produced 17.1 (*SD* = 5.39) unique labels per stimulus, and these labels indicated that participants perceived an average of 2.65 (*SD* = 0.06) levels of hierarchy per stimulus. Instruction format (top-down/bottom-up) did not have an effect on the mean number of hierarchy levels indicated by participants across stimuli (top-down instructions number of indicated levels: $M = 2.69$, $SEM = 0.15$; bottom-up instructions number of indicated levels: $M = 2.52$, $SEM = 0.10$; $t(29) = 0.94$, $p = .35$).

As the free response paradigm allowed for significant variability in the participant responses, the most frequently occurring grouping pattern for each stimulus was identified, and the degree of agreement between participants for these "modal" patterns was evaluated. It was observed that, on average, 24.1% of participants agreed upon the most commonly occurring grouping pattern for a given stimulus; however, there was disparity in this degree of agreement across stimuli, where consensus on the "modal" grouping pattern between participants ranged from 51.6% agreement to as little as 0.0645% (see Figure 3 for a visualization of degree of agreement across stimuli). Due to this variability in the degree of agreement between participants across stimuli, we evaluated whether the degree of agreement in grouping of a stimuli was related to the number of visual objects or features in the stimulus; however, degree of agreement on the most frequently occuring grouping pattern for each stimulus was not related to either the number of visual objects ($r = -0.26$, $p = .14$) or the number of visual features included in the stimulus ($r = -0.01$, $p = .95$).

As the majority of stimuli were designed to contain four specific structured relationships between the included visual features, we evaluated the modal grouping patterns of the 28 (out of 34 total) stimuli that included these relationships. For stimuli that contained both 20mm by 20mm squares and "nine-sets" comprised of the "small squares" (11 total stimuli), it was observed that the modal grouping patterns indicated individual "small squares" as residing on the lowest level of hierarchy, and squares as residing on an intermediate level of hierarchy, in all but one case (90.9%). In this remaining case, "small squares" were not individually identified by the modal grouping pattern, and squares were indicated as being on the lowest level of hierarchy; however, this stimulus had effectively no agreement between participants on a modal grouping pattern (see the value for "stimulus index 31" in Figure 3). In this same set of stimuli, participant's modal responses identified square visual objects and "nine-set" visual features as occurring at the same level of hierarchy for all stimuli. The relative position of these visual objects/features in the grouping hierarchy was intermediate for all but one stimulus, where they were grouped on the lowest level of hierarchy, however, this occurred on the same "disputed" stimulus.

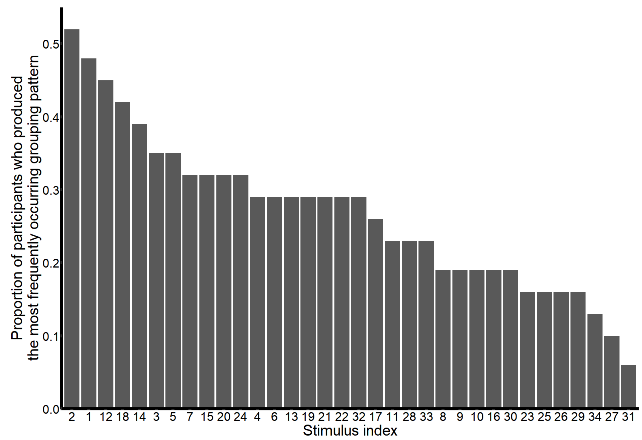

Figure 3: The proportion of participants who produced the most frequently occurring grouping pattern, across stimuli.

When a stimulus contained both square and rectangular visual objects (12 total stimuli), participant's modal responses identified square and rectangular visual features as occurring at the same level of hierarchy for only one stimulus (8.3%). Instead, participants were more likely to identify rectangles as being on the same level of grouping hierarchy as a row or column of three square visual features, with the square visual features being grouped independently at a lower level of hierarchy (11 of 12 stimuli, 91.7%). Similarly, when a stimulus contained both "nine-set" and rectangle visual objects/features (12 total stimuli), participant's modal responses never identified these visual features as occurring on the same level of hierarchy. Akin to relationships between "small squares," squares, and rectangles, the "nine-set" visual feature was always placed on an intermediate level of hierarchy, above the individual small squares that make up the "nine-set" and below the row/column structure associated with rectangles. Finally, in 100% of cases where all three kinds of visual feature were present in a stimulus (6 total stimuli), "small squares" were identified by participant's modal responses as being on the lowest level of hierarchy, square and "nine-set" visual features were identified as being on the same (intermediate) level of hierarchy, and rectangles were assigned to a more inclusive (higher) level of hierarchy containing row/column features made up of aligned square/"nine-set" features.

As the stimuli were designed to contain row and column structure, we evaluated the frequency with which participants identified this structure. Whole rows or columns were identified by the modal grouping patterns in 82.1% of cases across the 28 stimulus that possess the 3x3 grid design. When the stimulus contained horizontally aligned square and/or "nine-set" features and at least one horizontal rectangle, rows were identified as a grouping pattern in 100% of cases ($N = 9$ stimuli), while columns were identified as a grouping pattern in 0% of cases. When the stimulus contained vertically aligned square and/or "nine-set" features and at least one vertical rectangle, rows were identified as a grouping pattern in 0% of cases ($N = 9$ stimuli), while columns were identified as a grouping pattern in 88.9% of cases. For the stimuli in which rows or columns were not identified by the modal response (5 out of 28), visual inspection determined that the visual features of all but one of these stimuli were equidistant and did not contain at least one rectangle in either a horizontal or vertical orientation.

Finally, we were curious as to whether the grouping structure described by the modal responses ("small squares" positioned at a lower level of hierarchy, squares and "nine-sets" positioned at an intermediate level, and rectangles positioned at a relatively higher level) was also present in participant responses that did not adhere to the modal grouping pattern. For the set of stimuli that contained at least two of the three features (squares, rectangles, and "nine-sets" comprised of "small squares"), we evaluated non-modal participant responses that contained at least three levels of hierarchy, and categorized the relative level of hierarchy that these features were grouped on by participants.

It was observed that "small squares" were grouped at a low level of hierarchy in 100% of cases (216 total observations). Squares were grouped at the lowest level of hierarchy in 64.4% of cases, at an intermediate level in 34.6% of cases, and at a high level in 0.976% of cases (205 total observations). Comparatively, "nine-sets" appeared at a low level of hierarchy in 5.6% of non-modal responses, at an intermediate level 90.9% of the time, and at a high level in only 3.4% of cases (232 total observations). Rectangles were positioned at a low hierarchical level in 31.5% of non-modal participant responses, at an intermediate level in 64.3% of cases, and at a high level in 4.3% of responses (235 total observations). A chi-squared test indicated that, in these non-modal responses, there was a significant relationship between the type of indicated group ("small square", square, "nine-set", and rectangle) and the indicated level of hierarchy ($\chi^2(6, N = 888) = 450.47$, $p < .001$); a Fisher's exact test on the same contingency table also reported a significant relationship ($p < .001$). See Figure 4 for a visualization of the appearance of these labels across relative levels of hierarchy in non-modal participant responses to stimuli that contained at least two of the square, "nine-set," and rectangle visual features.

Discussion

In the present study, participants provided free-response groupings of visual scenes comprised of a set of visual features/objects with consistent structural relationships. Frequently occurring grouping patterns were evaluated to determine whether the structure contained within the task's stimuli was related to hierarchical structure of the participant's grouping responses. Singleton stimulus objects that differed in size but were identical in shape (4mm by 4mm "small squares" and 20mm by 20mm squares) were very likely to be distinguished by the modal responses as being on separate levels of a grouping hierarchy, with "small squares" consistently found on the lowest level and "squares" on an intermediate level. However, when a singleton visual object was compared to a visual feature which was identical in area but differed in the number of independent elements contained by that area (squares vs. "nine-sets"), all modal responses identified these stimulus elements as occurring on the same level of a grouping hierarchy. This influence of form over number of contained elements was also observed when the grouping of rectangular visual objects was compared to aligned square and/or "nine-set" features–in the majority of cases, horizontally/vertically aligned visual features that contained multiple independent visual objects were identified as being on the same level of grouping hierarchy as rectangles that were identical in area. Additionally, the set of non-modal responses to stimuli containing these visual features suggest the same general hierarchy observed in the modal responses–however, in these responses, "nine-sets" were distinguished as occurring on a higher level more frequently than squares.

Combined, these observations suggest that differences in size/areal span contribute more strongly to distinguishing objects and/or features with similar or identical shapes as being on different levels of a hierarchy, when compared to the singleton vs. multi-element nature of a feature. Visual features that contain upwards of 27 independent visual objects were found to be on the same level of grouping hierarchy as singleton objects; however, these sets of objects had to be positioned in an array that approximated the size and shape of the singleton object. This similarity is important–for example,

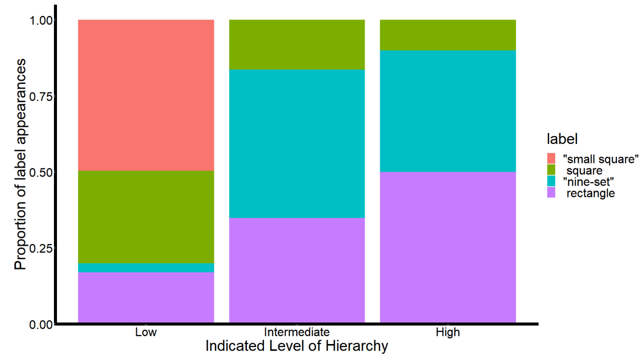

Figure 4: The proportion of "small square," square, "nine-set," and rectangle labelings at each level of hierarchy in non-"modal" responses that contained at least three levels of hierarchy.

no participants identified an "L" shaped set of three visual objects/visual features as being a group at all, let alone as a group on the same level as a rectangle. What remains unclear, however, is whether the presence of singleton objects that are similar in form to arrays of independent visual objects drive the grouping of those arrays, or whether the form of the array of visual objects itself is enough to drive grouping.

The insights gathered through this paradigm can contribute to the design of theoretical and computational models of visual grouping behavior by providing information on how object/feature characteristics such as size, form, and alignment are associated with perception of hierarchy in visual scene groupings. Previously, a visual grouping model that relies solely on proximity to predict of groupings of visual objects within a scene has been presented (Lindstedt & Byrne, 2018). While this model is capable of producing meaningful groupings, it cannot represent hierarchically grouped features; additionally, manipulation of the proximity of a scene's visual objects can result in improper grouping of scene that is easily parsed by humans. It is possible that proximity alone can be used to produce hierarchical groups; multiple radius thresholds can demarcate transitions between levels of hierarchy, or multiple "passes" of grouping can be performed to iteratively group the scene (where each "pass" of grouping produces a new level of hierarchy). However, the participant free response groupings in this study make it clear that objects/features in a visual scene are grouped according to features above and beyond proximity alone; in addition, consideration of additional visual features have the potential to make a model more robust to oddities in the proximity of the constituent visual features.

An effective model of human visual grouping processes would not only incorporate characteristics beyond proximity, but also make distinct predictions regarding how these features influence the structure of a hierarchically organized set of groupings. Here, it is apparent that the size of visual objects/features distinguishes the level of hierarchy that a particular visual object/feature is placed within–"small squares" and squares differed only in size, and were consistently placed on different levels of grouping hierarchy (with "small squares" occurring at a lower level than squares); a similar relationship was observed between squares and rectangles. Visual form, on the other hand, was more likely to distinguish separate groups within a level of hierarchy–for example, squares and "nine-sets" being identified as separate groups on the same level of hierarchy. A visual grouping model may determine the number of levels present in a given scene by first classifying each object into groups of similar size; the subsets of objects within each of these "size groups" could then be evaluated for form for further grouping.

However, further work is necessary to more fully delineate the influence of these features on visual grouping processes. For example, while the general trend of size and form was apparent in both modal responses and non-modal responses, some individual responses labeled objects of different sizes as residing on the same level of hierarchy; others marked "nine-sets" as being on a higher level than squares. In the first case, it is possible that these individuals relied on similarity of shape over similarity of size to differentiate levels of the hierarchy. Here, there were only two shapes (square and rectangle), and only two sizes of the square visual object, impeding sensitivity to this type of preference. In the second case, the "feature" nature seems to supersede the singleton square visual object. More careful and consistent variation of the relationships between the characteristics considered here would provide greater ability to determine the consistency of the influence of these characteristics on visual grouping behavior.

In conclusion, the results of the present study suggest that factors above and beyond proximity, such as size and form, contribute to human visual grouping processes. These factors are influential in how visual objects and features are distributed within a hierarchical set of groups. Understanding of the contribution of these factors to human visual grouping behavior can be utilized to improve computational models of human visual grouping behavior. In turn, these models can be applied in the design of interfaces to guide users towards proper understanding and intended behavior.

Acknowledgments

This research was supported by grant from the National Science Foundation. The views and conclusions contained herein are those of the authors and should not be interpreted as representing the official policies or endorsements, either expressed or implied, of NSF, the U.S. Government, or any other organization.

References

Byrne, M. D., & Kirlik, A. (2005). Using computational cognitive modeling to diagnose possible sources of aviation error. *International Journal of Aviation Psychology*, *15*, 135–155.

Greene, K. K. (2010). *Effects of multiple races and header highlighting on undervotes in the 2006 sarasota general election: A usability study and cognitive modeling assessment* (Doctoral dissertation). Department of Psychology, Rice University, Houston, TX, USA.

Lindstedt, J. K., & Byrne, M. D. (2018). Simple agglomerative visual grouping for ACT-R. *Proceedings of the 16th International Conference on Cognitive Modeling*, 68-73.

Norden, L., Kimball, D. C., Quesenbery, W., & Chen, M. C. (2008). *Better ballots*. New York, NY.

Salvucci, D. D. (2006). Modeling driver behavior in a cognitive architecture. *Human Factors*, *48*, 362–380.

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.