Perception & Psychophysics
2008, 70 (7), 1165-1184
doi: 10.3758/PP.70.7.1165

Set recognition as a window to perceptual and cognitive processes

Michal Jacob and Shaul Hochstein
Hebrew University, Jerusalem, Israel

The Set visual perception game is a fertile research platform that allows investigation of perception, with gradual processing culminating in a momentary recognition stage, in a context that can be endlessly repeated with novel displays. Performance of the Set game task is a play-off between perceptual and conceptual processes. The task is to detect (among the 12 displayed cards) a 3-card set, defined as containing cards that are either all similar or all different along each of four dimensions with three possible values. We found preference and reduced response times (RTs) for perceiving set similarity (rather than span) and for including cards sharing the most abundant value in the display, suggesting that these are searched preferentially (perhaps by mutual enhancement). RT decreases with number of sets in the display according to a horse race model, implying independence of simultaneous searches. Central cards are included slightly more often, but set card proximity seems irrelevant. A supplementary experiment determining dimensional salience showed consistent but individual preferences, yet these seemed not to affect set identification. Training induced gradual improvement, which generalized to a new version of the game, suggesting high-level learning. We conclude that elements of perception such as similarity detection are basic for finding sets in this task, as in other real-world perceptual and cognitive tasks, suggesting the presence of basic similarity-perceiving mechanisms. The findings confirm the conclusion that conceptual processes are affected by perception.

Often, a complex task is challenging, cognitive, analytical and effort demanding at first but later, with extensive training, becomes more immediate and perceptual and does not require focused attention and conscious processing. One example of such a process that has received an enormous degree of study is categorization (Ashby & Maddox, 2005). Another such process is reading, where the original, almost painful working out of a word or sentence becomes automatic (Carr, 1992)—even when contraindicated, such as in the Stroop effect (Stroop, 1935). The transition between these modes of processing requires further study. A related issue that has not received much attention is the visual-perceptual influences on these processes—when they are still cognitive, conscious, and analytic. We chose as the substrate for our study the game called *Set*. For the novice, it requires slow cognitive and conscious analysis of the display. Although these processes gradually become easier, quicker, and more perceptual, here we study conscious and unconscious perceptual influences on performance while the game is in its cognitive stage. Our ultimate goal is to understand better the perceptual mechanisms underlying set recognition as an instance of perceptual processes influencing cognitive tasks in general.

The Set *visual perception* game, demonstrated in Figure 1A, was invented in the U.S. by Marsha Jean Falco in 1974 (Set Enterprises Inc.; homepage: www.setgame.com). In the original game, each card has four *dimensions*, or attributes, and one out of three possible *values* for each, as follows: shape (diamond, ellipse, or wave), number (one, two, or three), filling (empty, striped, or filled), and color (originally, red, purple, and green; in our version, red, blue, and yellow). On each round, 12 cards are displayed on the table or, in our case, on the computer screen.

The goal is to identify a set, defined as 3 cards, being *all different* or *all alike* within *each* dimension, *independently* of the other dimensions. That is, along *each and every* dimension, the set either *spans* all values or has only one value—*similarity* within that dimension. In general, a valid set will span some dimensions and have similarity for others. Note that for any 2 cards, there is exactly 1 card that completes the set. As an example, in Figure 1A, the rectangles around 3 cards identify a set. There are two other sets present in these 12 cards. Can you find them?

Is playing the Set game a perceptual or a cognitive task? At first thought, since the game depends on viewing colorful geometric elements, it might be thought to be a perceptual task. In fact, the commercial version is called "The Family Game of Visual Perception." But this is a naive view, since the determination of whether cards form a set is a conceptual matter. Thus, the real question is whether the task is purely cognitive or whether there is also a perceptual element (other than the trivial aspect that we need to perceive the cards before we can perform the cognitive processing required to determine which constitute a set). That is, do perceptual characteristics influence the conceptual detection of a set (Goldstone & Barsalou, 1998)?

M. Jacob, michaljb@gmail.com

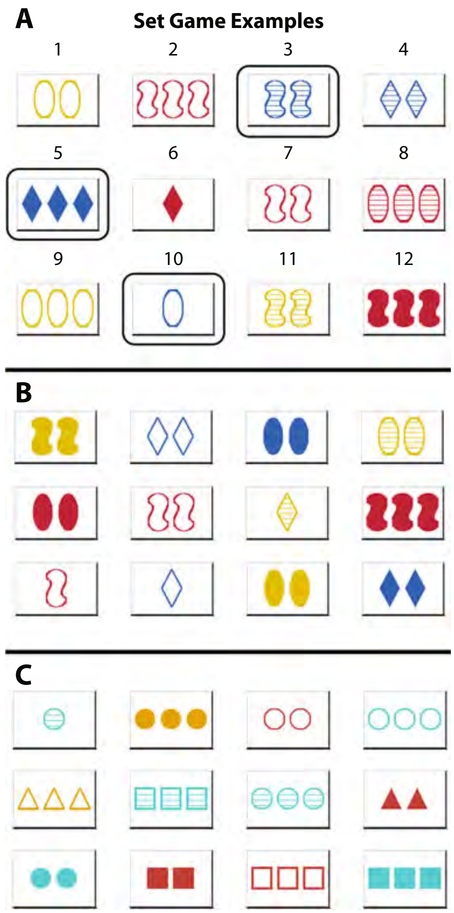

Figure 1. (A) The four-dimensional three-value Set game. The goal is to identify a set, defined as three cards, all of which are different (span) or alike within each dimension, independently of the other dimensions. Class is defined as the number of dimensions spanned within a set. The marked cards form a set of class 3; that is, they are all blue and span the dimensions of shape, number, and filling. There are two more sets in the display. Can you find them? The numbers above the cards are for identifying their location in the article; they did not appear in the actual task. The most abundant values (MAVs) in this display are red, wave, two, three, and empty, all with a group size (MAV-GS) of five. (B) Display demonstrating the four classes. Class 1 Cards 3, 5, and 11 span the colors, but all contain two (number) filled (filling) ellipses (shape). Class 2 Cards 1, 5, and 12 span the colors and shapes, but all contain two filled items. Class 3 Cards 4, 6, and 12 span the colors, shapes, and filling, but all contain two items. Class 4 Cards 4, 8, and 10 span all four dimensions (i.e., the cards contain exactly one each of all the possible values of each dimension). In this display, the MAV is the number two (cards with two items), with a MAV-GS of eight. (Note that in all, nine cards are included in the four sets, due to overlapping sets—i.e., cards belonging to more than one set.) (C) Generalization version of the game with new values. Sets in the display: Cards 1, 4, and 9 (class 2), 4, 5, and 11 (class 2), 1, 2, and 3 (class 3), and 1, 5, and 10 (class 4). If one of the first two sets seems to you to contain more similarity than the other, this may hint at your individual dimensional salience (see below, Experiment 2). The most abundant value here is three, with a group size of 7. If turquoise (6) or circle (6) seems to you more abundant than the number three, this may hint that the color or shape dimension is more salient than the number dimension. The additional sets in panel A are Cards 6, 7, and 8 (class 3) and 3, 6, and 9 (class 4).

As has already been pointed out, the Set game depends on detection of a mixture or combination of similarity and span. In the dichotomy or continuum between conceptual and perceptual processes, it would seem that similarity is more perceptual and spanning is more conceptual. That is, as the Gestalt psychologists pointed out, similar elements are naturally and quite automatically grouped together perceptually. One of the questions that we address here is whether there is an equivalent mechanism that automatically groups elements that span a dimension.

Another aspect of the Set game is its relationship to the general task of categorization. Categorization is the relating of objects or elements that differ in *irrelevant* ways, because they are similar in those features that are deemed *relevant*. Rosch and Mervis (1975) based their very definition of basic-level categories on maximizing similarity among category members, together with maximal dissimilarity between members of different categories, but some differences must remain among members of the same category or they would be identical. Bower and Trabasso (1963) analyzed performance of a concept identification task—where, again, one or more dimensions are relevant and others are present but irrelevant—and showed that in this rule-based task, subjects show all-or-none learning of potential rules (i.e., they “try” one rule at a time, sequentially testing the relevance of each dimension), rather than using all the information presented to them (by experimenter feedback).

Another well-studied example is the Wisconsin card sorting task (WCST; Berg, 1948; Grant & Berg, 1948; Heaton, Chelune, Talley, Kay, & Curtiss, 1993), a very simple categorization task in that only one dimension (of three possible) is considered *relevant* at any time (with the subject’s task being to find that dimension as it changes occasionally). In the Set game, too, cards must be associated on the basis of their similarity in some dimensions, despite their dissimilarity in others. The tasks are also similar in that they have common dimensions, with several possible values for each: The WCST has three dimensions (color, shape, and number, with four values along each dimension); the Set game has these same dimensions plus the additional dimension of filling (bringing the total to four dimensions, but here with three values along each). The Set game includes an added complication in that the number of dimensions for which there is similarity is *not* announced in advance and may change from display to display. But variability is inherent in the WCST, too, since the choice of which dimension is relevant changes from time to time without prior notification. For both tasks, there is a binary rule for each dimension and for each trial: Subjects decide which dimension is currently relevant versus irrelevant to the WCST (only one is relevant at each time) or which reflects similarity versus span for the Set game. Thus, in both cases, the subject must be ready for change and must not stick to old habits; both tasks require flexibly adjusting which cues are important in the environment. However, in the Set game, feedback is not essential for playing correctly, whereas in the WCST, the task relies on feedback. Another important difference between the WCST and the Set game is that the goal in the first is to match a card (a sorting task) and the goal in the second is to detect three cards according to the given rule. The most important aspect of similarity in the two tasks, however, may be the one described in the preceding paragraph—namely, that both are inherently conceptual tasks with a large degree of perceptual influence.

Thus, the Set game may be seen as a categorization task in that subjects have to find three cards that belong together, with similarity and dissimilarity along different dimensions. But Set qualification requires adherence to another rule that is added to the usual similarity requirement: In every dimension for which there is not full similarity, there must be full spanning; that is, the set cards must all be the same—or all different—for each and every dimension. In no case may there be two cards that share one value (say, red) and a third card that differs from them (say, blue). Set recognition is a difficult task because the number of *relevant* dimensions for which similarity must be found is unknown (and may even be zero) and the other dimensions are not simply *irrelevant* but need to span the possible values. Adding this simple requirement changes the nature of the task, and the ways in which it does so is the theme of this article.

In regard to detection of similarity versus spanning along the different dimensions, we introduce the term *class* as the number of dimensions spanned within a set. To elaborate, there are different types of sets, having different numbers of dimensions imposing similarity. We regard the complementary number, the number of dimensions fulfilling the criterion of difference (i.e., *span*ning all the values along that dimension), as the class of the set. Therefore, the class of a set represents the degree of dissimilarity within the set. For example, in Figure 1A, the marked set belongs to class 3, as does also the set of Cards 6, 7, and 8; the remaining set (Cards 3, 6, and 9) is of class 4. In Figure 1B we present a display that is designed especially to include one set of each class (see the figure caption). One of our goals is to determine whether players have a preference for sets of lower or higher class—that is, whether the ease of detecting a set depends on its degree of similarity (Tversky, 1977).

For a general game of n dimensions, there are n different classes, 1 . . . n. Thus, in the original game, there are four classes, numbered 1–4. There are no sets of class zero, because, by definition, this would mean that they are similar on all dimensions, or identical, which is not possible, since there are no repeat cards in a pack.

Various strategies may be used to find sets (see Holyoak, 1990, on problem solving). There are the obvious, exhaustive search strategies, of choosing each possible pair of cards and checking whether a complementary third card is present in the display or choosing each possible group of three cards and checking whether they form a set. These strategies require 220 ($= \binom{12}{3}$) operations, which a computer program does most easily but is not realistic for humans (see an alternative strategy in Box 1).

Obviously, there must be other strategies as well. Thus, although the task is easy from an algorithmic point of view and is readily solvable by a computer program, the exhaustive search strategy is not the way that the human brain computes and reaches a solution (see Ullman, 1984). Ultimately, the best strategy may be not to use a strategy but to catch the gist of the scene (Hochstein & Ahissar, 2002; Oliva & Torralba, 2006; Schyns & Oliva, 1994) and preattentively let the sets pop out. Thus, the Set game provides a window to a "solved" problem, one with a creative solution, that is, for the time being, beyond our comprehension. Our approach includes psychophysical experiments and combinatorial analysis. In Box 1, we present general combinatorics for the Set game.

In Experiment 1, the subjects played the game, and we recorded their choice of cards and timing of performance. We found that the subjects preferred sets of lower class, and we analyzed set perceptual parameters that affected set cognitive search strategy, which set was detected (when more than one was present), and the speed of its being detected, in terms of response time (RT). These included number of sets present in the display, abundance of different values, place effects, and influence of previous card locations.

In Experiment 2, we used another experimental paradigm to determine the relative salience of different dimensions present in the Set game—in a subject-by-subject manner. The results were compared with the prevalence of these dimensions or values in sets found by the same subjects when they played the Set game.

In Experiment 3, we tested the dynamics of learning effects following experience with the game. Training improved performance dramatically in terms of speed of detecting sets. We also tested the effect of training with one version of the game on performance with a different version (new values along the original dimensions; see Figure 1C)—that is, the degree of learning generalization and transfer of learning effects. That is, after training has improved performance, will changing stimulus values start learning all over again, or is the ability to identify a set now established, involving higher cortical level mechanisms regardless of specific lower level stimuli (Ahissar & Hochstein, 1997, 2004; Hochstein & Ahissar, 2002), so that search with new values is performed immediately at the lower, posttraining RT?

EXPERIMENT 1
Playing the Set Game

Method

We implemented the Set game with an interactive computer program, allowing us to record subject moves and RTs. On every round of the game, 12 cards were displayed, always including at least one set. The subjects were instructed to mark (via mouse clicks) three cards that formed a set as quickly as possible but to try to avoid mistakes. If the three chosen cards indeed formed a set, a Continue button appeared. Upon pressing of the Continue button, three new cards were dealt in place of the three marked set cards, and a new round began. The replacement cards were taken from the remaining "deck," with the used set cards being excluded. During the round, if a player changed his or her mind after marking fewer than three cards, the player could "unmark" the cards by a second mouse click. In case the player chose three cards that did not form a set, a "try again" message appeared, along with an explanation of why the chosen cards did not form a set (e.g., "Two are blue and one is red"). If 10 such mistakes occurred, the set was revealed, and the round was counted as unsuccessful. RT was measured from the pressing of the Continue button until the third card of a set was chosen; the subjects were informed of this timing procedure. This process continued, going from round to round, until the entire deck had been used, com-

BOX 1
Combinatorics: General

Combinatorial analysis of the Set game is presented in the relevant parts of the article, revealing a number of game characteristics and explaining individual performance heuristics by comparing behavior with task parameter distributions. Here, we will give only a preface to the combinatorics.

We denote the number of dimensions as d, and the number of values along each dimension as v. Note that v is the size of a set, and d may be considered the complexity of the game.

The total number of (different) cards is $v^d = 3^4 = 81$ in the regular case of a four-dimensional three-valued game. Each combination of two cards uniquely defines a third card for a set, and there is a third card that constitutes the set for any two: There is one and only one card completing a set. This can be elaborated by a vector representation, as follows: Each card is represented by a four-dimensional vector, and in each dimension it receives one of the three possible values. Two cards, for instance, can be [3, 1, 2, 1] and [2, 1, 2, 3]. The vector of the third card of the set is constructed by the following rule: If the two cards have the same value for a certain dimension, use this value also for the third card; if the two cards have different values, use the complementary value. This will always lead to a unique and existent card. In the case of the two cards mentioned above, the third is [1, 1, 2, 2]. The number of all possible sets is $(81 \cdot 80)/3! = 1{,}080$, since there are 81 possibilities for choosing the first card, one less (80) for choosing the second, and only one way of completing the set. Division by (3!) excludes repetitions.

Number of Triplets in the Display

As was mentioned above, the number of possible choices of 3 cards out of 12 is $\binom{12}{3} = 220$. An alternative calculation, choosing 2 cards at a time and verifying that the complementary 3rd is present sums to the same number of operations as follows. Choose the first 2 cards and verify whether any of the other 10 cards will complete a set (10 operations). Then, in the inner loop, increment the 2nd card to the 3rd in the display (thus now choosing the 1st and 3rd), leaving 9 cards for verification. Continuing the inner loop results in a decreasing arithmetic series (10 . . . 1) of operations. Now for the outer loop, increment the 1st card to the 2nd in the display, repeating again the inner loop, this time from the 3rd card, giving a decreasing arithmetic series (9 . . . 1), and so on. Thus, the entire process yields $\sum_{m=1}^{10}\sum_{n=1}^{m} n = 220$ operations.

An Alternative Strategy for Finding a Set

In the introduction, we referred to possible exhaustive search strategies for finding sets. An alternative strategy is called *dimension reduction*. The idea is to choose a dimension (preferably the one that is perceived as most dominant for the player, or one with the most abundant value). On this dimension, denote the three values as x, y, and z. Look at all the cards along this dimension that have the value x, and among them search for a set. This dimension is thus set to be fixed, reducing the number of dimensions by one; that is, the set will have similarity on this dimension with the value x assigned to it. If no such set exists, move on to y and then to z. If, in all cases, no set is found, the conclusion is that the set must span this dimension. So choose the next salient dimension and repeat the process, now looking always for cards spanning the previously chosen dimension(s). If the game is four-dimensional, it is usually enough to reduce one dimension (no need to set two dimensions at a time to be fixed), because it is perceptually quite easy to find a set in three dimensions, especially within the smaller group of cards. This strategy is especially relevant to the impact of the most abundant value, described below.

pleting a *game*. Note that in our version of the game (but not in the original), we never presented a display without any set.

Before beginning to play the game, the subjects read an explanation of the task and the definition of a set. They were shown a sample display with one set marked (as in Figure 1) and were asked to find another set in the display (as you may do in Figure 1). Incorrect responses were explained, so that the subjects would understand the nature of a set and that a set might not contain two cards that were similar along any one of the four dimensions (e.g., two cards with striped elements) and only one card that was different on this dimension (e.g., one card that had filled elements). The experimenter verified that they understood the rule before starting the first session.

The subjects participated in several sessions, each lasting about an hour, in which they played several games, as time permitted (according to their individual performance level). Each game included up to 24 rounds (the maximal number in an 81-card pack, displaying 12 cards and replacing 3 each round). Sometimes there were fewer rounds, to avoid occurrences of displays without a set. Eleven subjects participated in three sessions, each playing 3–6 games per session (following the first, slower session); another 11 participated in fewer sessions (9 in one session with 1 or 2 games, 1 in a session with 3 games, and 1 in two sessions). In total, 22 subjects played 125 games, including 2,844 rounds—that is, displays and sets chosen. We eliminated the 1st round in each game because of its exceptional conditions (the subjects had not previously seen any of the cards, whereas on subsequent rounds only 3 cards were new). We also disregarded outlier rounds with exceptionally long RTs (setting a bound at the mean RT plus three times the global standard deviation); 2,664 rounds remained. In 9% of the 2,844 rounds played, the subjects marked 3 wrong cards and were informed they did not form a set; in 11%, they marked 2 cards and then "unmarked" them (including overlapping rounds).

MATLAB was used for experimental control, data collection, and analysis. We measured RT from display presentation to choice of the third card completing a set, tracking cards, the detected set, and other available sets and their positions, deriving the dependence of preference on the following parameters: similarity (set class), number of sets present in the display, the most abundant value present, the distance of set cards from each other and their location in the display, and the relationship of cards chosen to cards just revealed (i.e., locations used in the preceding set).

Results and Discussion

Similarity

As was mentioned in the introduction, we were interested in the impact of the class of a set on the speed with which it was detected, as well as the choice of it over sets of other classes when more than one class was present. Figure 2A shows mean RT ($\pm SE$) by class. On average, the lower the class, the shorter were the RTs. Figure 2B (black bars) shows the number of sets detected from each class. Clearly, more sets were detected for higher class sets (up to class 3). Does this trend reflect a true preference on the basis of set class, or does it simply reflect their abundance? To answer this question, we compute how many sets of each class are available in Box 2.

Consider the results of Figure 2B, looking at the sets detected from each class, but now taking into consideration the relationships among the number of sets available of each class. Only twice as many sets were detected from class 2 as from class 1, even though combinatorially there are 3 times as many sets from class 2. Class 3 sets were detected just slightly more often than those of class 2, although there are 1.33 times as many such sets of class 2 (4 times the number of class 1 sets). Class 4 sets were detected slightly less often than those of class 1, although there are twice as many such sets. Thus, there seems to have been a preference for lower class sets. This is also demonstrated in Figure 2B, in which we show the empirical distribution of rounds (gray bars) and occurrences (white bars) of sets of different classes, and in Figure 2C, in which we show the number of detected sets as a fraction of these distributions.

Another way of testing preference for sets of different classes is to analyze the relative place of chosen sets among available sets on that round. We plot the probabilities that a set would be chosen when there were other sets present on that round of either lower (Figure 2D) or higher (Figure 2E) class (the abscissa reflects the number of sets present on each round from classes that were lower or higher than the one chosen, and the ordinate is the number of occurrences of such choices, normalized to the number of opportunities for such a choice—i.e., the number of rounds on which it was possible to choose a set and leave that number of sets of lower or higher class). Note that Figure 2E is flat, reflecting an equal probability of leaving any number of higher class sets: The players were indifferent to the presence of higher class sets. In contrast, Figure 2D declines with the number of lower class sets. Taken together, these graphs demonstrate a preference for lower class sets—that is, for sets with more similarity. We conclude that sets are found on the basis of a similarity-detecting process—presumably, a basic perceptual mechanism.

Recent results in the realm of categorization also point to the use of underlying mechanisms that perceive simi-

BOX 2
Combinatorics: Division Into Classes

The general equation for the total number of sets of class i, with d dimensions and three values, is

$$\frac{3^d \cdot \binom{d}{i} \cdot 2^i}{3!};$$

there are 3^d possible choices of the first card, $\binom{d}{i}$ possibilities to choose which i dimensions to change for the second card, and 2^i variations of changing the i dimensions. The third card is determined uniquely by the first two, and 3! again excludes repetitions.

Summing over i, this expression gives the total number of sets, $[3^d \cdot (3^d - 1)]/3!$. In particular, for $d = 4$,

Number of class 1 sets: $\dfrac{81 \cdot \binom{4}{1} \cdot 2^1}{3!} = 108 \equiv x;$

Number of class 2 sets: $\dfrac{81 \cdot \binom{4}{2} \cdot 2^2}{3!} = 324 = 3x;$

Number of class 3 sets: $\dfrac{81 \cdot \binom{4}{3} \cdot 2^3}{3!} = 432 = 4x;$

Number of class 4 sets: $\dfrac{81 \cdot \binom{4}{4} \cdot 2^4}{3!} = 216 = 2x.$

Note that these numbers sum to $1{,}080 = (81 \cdot 80)/3!$.

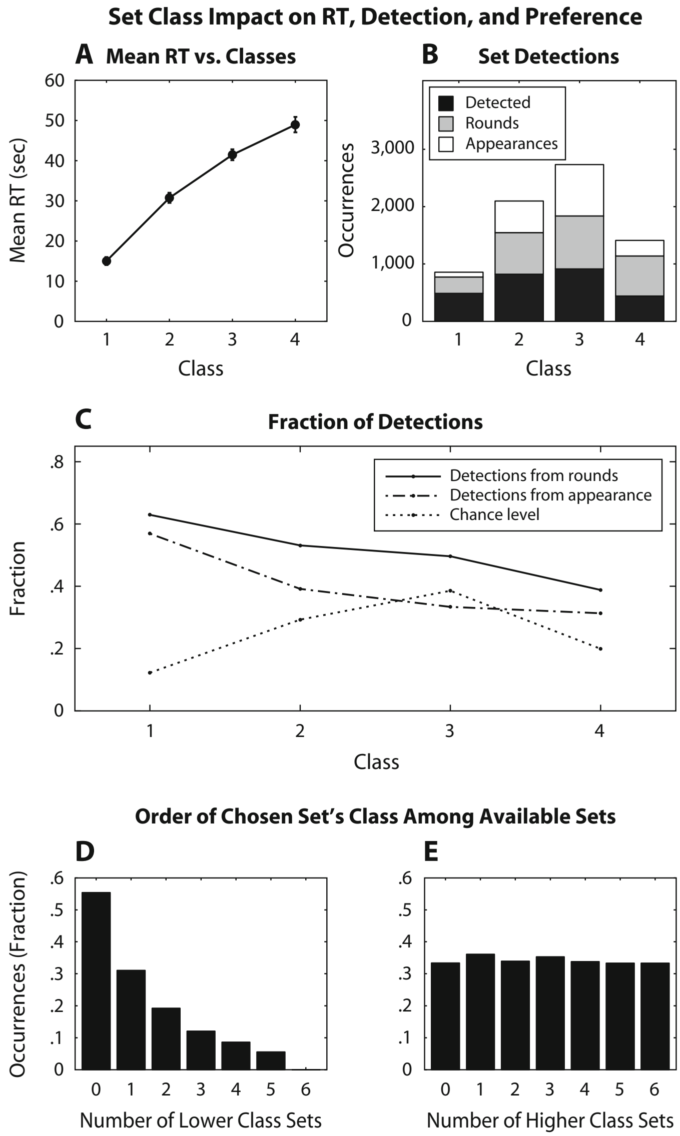

Figure 2. Influence of similarity on set detection, with data for 22 subjects, 125 games, and 2,664 rounds: The effect of the class of a set (number of dimensions with span, rather than similarity) on detection speed, as well as on its preference over sets of other classes when more than one set of different classes was present. (A) Mean response times (RTs) for detecting a set by class number; error bars here and in the other figures are the standard errors of the mean (*SEM*s). (B) Distribution by class of detected sets (black bars), number of rounds in which such sets appeared (gray), and total number of appearances (white), including possibility of more than one set in a display. (C) Set detections by class as a fraction of rounds (solid curve; black/gray bars in panel B) and as a fraction of total appearances (dashed curve; black/white bars in panel B), both showing a decrease as the class increases. For comparison, we show the chance probability (dotted curve) of choosing each class (which follows the pattern calculated by the combinatorics in Box 1). (D and E) Order of chosen sets among all available sets, demonstrating preference for sets of lower classes. Only (1,866) cases with at least two sets from different classes were included. (D) Number of choices of a set as a function of the number of lower class sets present on that round, normalized to the number of opportunities for such a choice. (E) Normalized number of choices as a function of higher class sets present. The decreasing distribution for lower class sets and the independence of higher class sets support a preference for finding lower class sets.

larity. It was found that the learning of categories is easier and more natural when one learns from exemplar pairs that belong to the same category than when one learns from pairs that belong to different categories (Hammer, Hertz, Hochstein, & Weinshall, 2005, 2007, in press), even when the pairs are preselected to contain the same amount of information. In addition, children are even more biased toward learning from same-class pairs (Hammer, Diesendruck, Weinshall, & Hochstein, 2008). Similarity is also the basis of Gestalt principles of grouping (Koffka, 1935; Köhler, 1929). Although the Set game is not a usual categorization task or a standard grouping phenomenon, it is interesting to speculate that the same basic mechanisms may underlie all these processes. In the introduction, we compared Set detection with the WCST and the concept identification task. Despite the differences mentioned there, the many similarities between these tasks in terms of the perceptual dimensions used and, more significantly, the need to choose which dimensions are relevant for any particular trial and to repress this choice for the following trial suggest that the same similarity-detecting mechanism may play an important role in all these tasks. This congruence of findings suggests that despite the cognitive nature of the tasks, perceptual mechanisms play an important and essential role.

Finally, it is worthwhile noting that the preference for similarity was not learned during the game, because, in the game itself, there was actually a greater abundance of higher class sets (see Figure 2B). Thus, the bias toward detecting lower class sets must reflect an innate or prior preference, perhaps stemming from the tendency for real-world categories to be organized around similarities, rather than around differences.

In summary, we found that sets from lower classes were detected more quickly and recognized more often (relative to their availability)—that is, with priority when more than one set was present—suggesting that greater similarity is a factor in set detection.

RT by Number of Sets Present

An obvious parameter that might have influenced the speed of detecting a set is the number of sets simultaneously present in the display, although the subjects were not informed of this number. Calculation of the expected distribution of number of sets is nontrivial, so we determined the empirical distribution in our experiment, as displayed in Figure 3A. Recall that in our experiment, we rejected displays without sets, although the original game does not.

Horse race model (theoretical analysis). The horse race model (Miller, 1982, 1986; Raab, 1962; Townsend & Ashby, 1983) deals with processes that compete with each other. When there is more than one stimulus competing for our attention, the result can be performance facilitation according to probability summation (Graham, 1989; Monnier, 2006), resulting from the processes being independent (Corballis, 1998; Monnier, 2006). Another term for that same phenomenon is redundancy gain (Corballis, 1998; Egeth & Mordkoff, 1991; Garner & Lee, 1962). Alternatively, on the one hand, if the processes give rise to synergistic neural summation (rather than their being independent), the result will be an *enhanced redundancy gain*, which is greater than simply probability summation. On the other hand, there can be interference between the processes. These properties will be true both for performance accuracy (Graham, 1989) and for RT (Raab, 1962). We extend the horse race model to N processes and apply this model to set detection; thus, we view the different sets present in a display as competing with each other to reach detection.

We wish to derive a theoretical graph describing mean RT versus the number of sets present in the display, as expected by probability summation—that is, the horse race model—and to compare it with the empirically measured RT dependence. As a first step, we measure the empirical (binned) distribution of RTs, $p(t)$, for detecting a set when only one set is present in the display. This is shown in the upper left curve of Figure 3B. The horse race model procedure is then to assume that the time for detecting either of the sets when there are two available is the minimum of the times for detecting each. Thus, the new RT distribution for two sets is formed by taking a pair of time bins, t_1 and t_2, having individual probabilities, $p(t_1)$ and $p(t_2)$, with a joint probability of $p(t_1) \cdot p(t_2)$, and assigning it to the time bin of the minimum of t_1 and t_2. The mean RT according to the horse race model, when there are two sets, will then be

$$\sum_{t_1=0}^{\infty} \sum_{t_2=0}^{\infty} p(t_1) \cdot p(t_2) \cdot \min(t_1, t_2).$$

By induction, the same procedure is used to derive the RT distribution when there are three or more sets present and, from these, the expected mean RT for each number of sets. If the processes are indeed independent, the observed RT distribution will fit this theoretically created distribution; if there is neural summation and, therefore, enhanced redundancy gain, there would be a shift to the left in the observed distribution (shorter RTs); if there is interference, there would be a shift to the right (slower responses).

Experimental results are compared with predictions of the horse race model in Figures 3B and 3C. We find a good fit between the empirical results and the model predictions, suggesting that the model can account for our results. The implication of the success of the horse race model is independence (rather than synergy or interference) of the processes of detecting each set when there is more than one present.

Most Abundant Value

The relative abundance of the values along each dimension within the array of 12 cards presented on each round may influence which set is chosen and the RT for finding it. We are particularly interested in the cards belonging to the largest of these groups of values, which we call the *most abundant value* (MAV), and the number of cards in this group, called the *MAV group size* (MAV-GS). Since similarity was shown to be easier to perceive, it may be that subjects search among the cards that belong to the largest group with the same value along one dimension—that is, the cards that share the MAV.

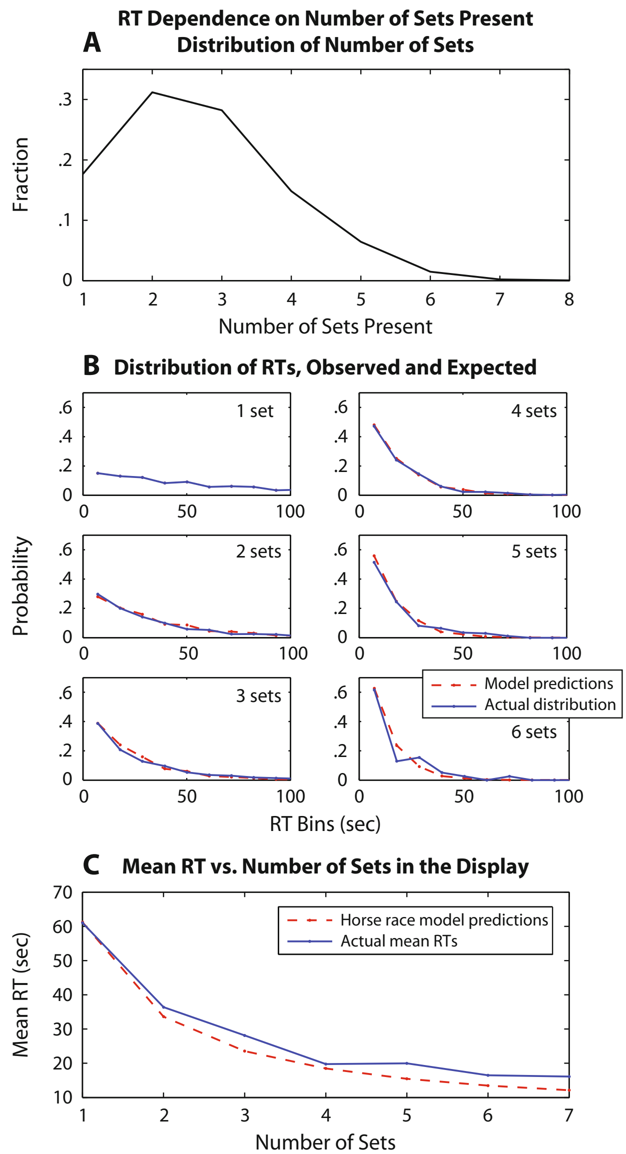

Figure 3. Influence of number of sets simultaneously present in the display, comparing experimental results with predictions of the horse race model. (A) Empirical distribution of the number of sets in the display (for the 2,664 rounds; recall that in the experiment, displays contained one or more sets). (B) Observed (solid blue) and expected (dashed red) response time (RT) distributions according to the horse race model for one to six sets present. As a first step, we measure the empirical distribution of RTs for detecting a set when only one set is present in the display (upper left curve). From this, we derive the expected distribution of RTs for two sets (see the text). By induction, the same procedure is used to derive the expected RT distribution when there are three or more sets present. (C) Mean RT by number of sets in the display, observed and expected. Compare the experimental results with the horse race model predictions—that is, performance facilitation according to probability summation, indicating independent processes. The alternatives, neural summation (enhanced redundancy gain) or interference, would have resulted in a shift of the graphs to the left or the right, respectively. The results show neither of these effects but, rather, a good fit to the model predictions, suggesting independence of search for different sets.

For example, in Figure 1B, the MAV is the number two (i.e., cards with two items) with a MAV-GS of eight (i.e., there are eight cards with two items; Cards 3, 5, and 11, Cards 1, 5, and 12, and Cards 4, 6, and 12 form sets from within the MAV group; Cards 4, 8, and 10 form a set outside the MAV group); in Figure 1C the MAV is the number three, with a MAV-GS of seven (Cards 4, 5, and 11 form a set from within the MAV group; Cards 1, 2, and 3, Cards 1, 4, and 9, and Cards 1, 5, and 10 form sets outside the MAV). In Figure 1A, the MAV-GS is five, and there are several groups with this size sharing some value: red (Cards 6, 7, and 8 form a set within this MAV group), wave, two, three, and empty.

Reviewing these terms, MAV refers to the most abundant *value* itself, a MAV *group* is the group of cards with that (most abundant) value, and MAV-GS is the actual *size* of the group.

In Figure 4, we demonstrate the stages in determining whether sets among the most abundant cards are preferred. In Box 3 we derive the theoretical MAV-GS distribution shown in Figure 4A (solid curve). Figure 4B shows the empirical MAV-GS distribution (dotted curve; in accord with the computed distribution in Figure 4A). The most common MAV groups are of six and seven cards. Within this MAV-GS distribution, those groups that included sets (2,251 of 2,664 rounds; 84.5%) are distributed as plotted in Figure 4B (dashed curve). The distribution of MAV-GS within which a set was actually detected by the player (1,619 of 2,251 rounds; 72%) is shown in Figure 4B (solid curve).

Now we average the fraction of rounds for each game in which the subjects detected a set among the MAV cards, out of the number of rounds in which such a set was present (Figure 4C, solid curve; error bars indicate standard errors). An almost monotonic increase is observed. That is, there was a trend for the number of discovered sets to increase within the MAV group as its size increased. This finding seems to suggest that the presence of more cards in the MAV group leads to a better probability of detecting a set within it. However, this may be misleading, since we should consider only cases in which there was a choice. Perhaps when the MAV group is large, that value is so dominant that there is no set without similarity in it. In this spirit, we add the corresponding chance level of finding a

BOX 3
Combinatorics: Distribution of Most Abundant Value Group Size (MAV-GS)

In any dimension, the cards on display may include one, two, or all three values, so that the group of cards with a particular value of a particular dimension may include anywhere from 0 to 12 cards. Calculation of the combinatorial statistics of occurrences of each MAV-GS is done according to the following equation, stating the probability of having three groups of sizes k, l, and $(n - k - l)$ out of n cards:

$$\frac{n!}{k!\,l!\,(n-k-l)!}\left(\frac{1}{3}\right)^n .$$

This is a probability, so that the sum of all possible distributions is 1. To satisfy this constraint, the first part of the following function (the factorials) must sum to 3^n. For example, for $n = 12$, the sum must be $3^{12} = 531{,}441$, as it is.

$$\sum_{k=0}^{n}\sum_{l=0}^{n-k}\frac{n!}{k!\cdot l!\cdot (n-k-l)!}\cdot\left(\frac{1}{3}\right)^n .$$

We construct all possible series of three groups in the display, with the total number in the three groups summing to n, the number of cards in the display; for example, for 12 cards, the following series would exist: (12, 0, 0), (11, 1, 0), (10, 2, 0), (10, 1, 1), and so on. Each series can appear, in a permutation, one, three, or six times (when there are three, two, or no repeated values, out of three, respectively); for example, the series above would appear three, six, six, and three times, respectively. The probability of each series is calculated by the equation above (with n being 12, and k and l representing two out of the three group sizes), multiplied by the number of permutations. Taking the largest value in each series and summing the probability of the related series yields the probability of each MAV-GS (the number of cards in the most abundant group), but only in one dimension (Figure 4A, dashed curve).

To calculate the probability for each MAV-GS when there are four dimensions, we construct a series of largest values for each dimension. Then we calculate the probability of each such series, according to the previously calculated probability of each value, and accumulate this to the probability of the largest value in the series (similar to the calculation for the shortest RT in the horse race model). Now we have the probability vector of each MAV-GS for four dimensions (Figure 4A, solid curve). The graph shows a shift to the right when the number of dimensions is raised, because the probability of having a low value as the most abundant in all dimensions becomes less likely the more dimensions there are. Each probability vector of course sums to 1.

A group of three of the same value cannot be the largest such group, because then there will be a group of at least $(12 - 3)/2 = 4.5$, meaning of at least five. Although the most abundant value can be a group of four cards, because the cards in a certain dimension can be divided into 4–4–4 groups of values, the probability of this occurring actually approaches zero, since this 4–4–4 division would have to apply to all four dimensions.

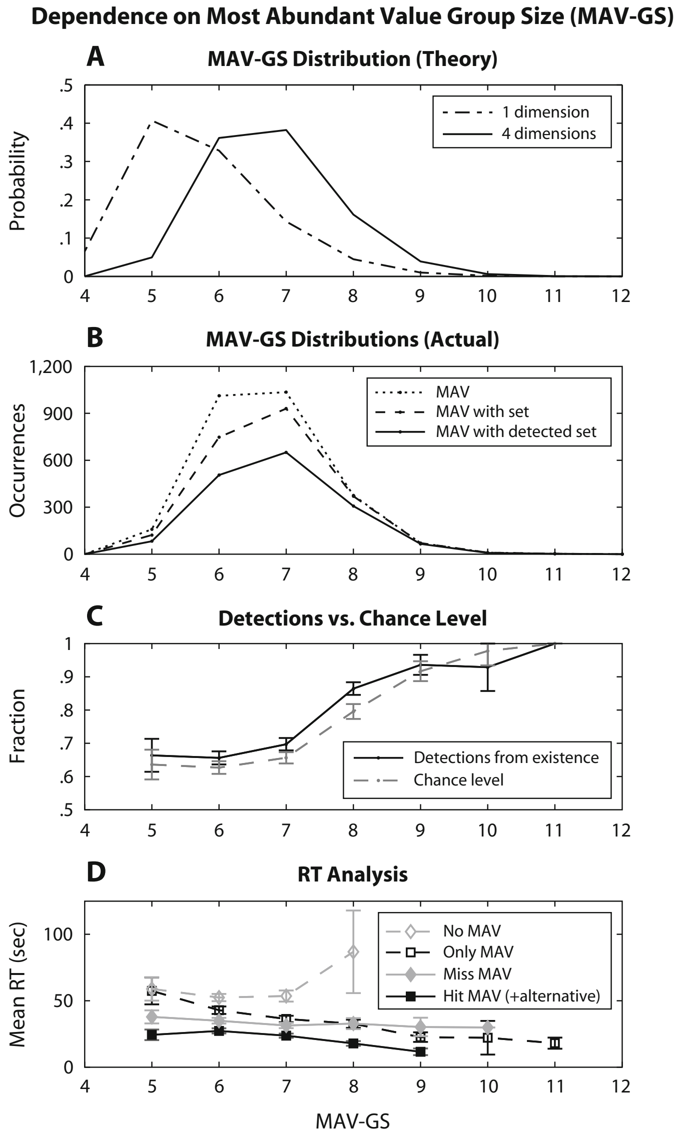

Figure 4. Most abundant values (MAVs). (A) Theoretical distribution of MAV group size (MAV-GS) of values, in one dimension only (dashed line) and in four dimensions (solid line); see the text (Box 3) for derivation. With more dimensions present, there are more chances for larger group sizes to appear. (B) Empirical MAV-GS distribution (dotted curve, similar to the theoretical distribution in panel A), distribution for largest groups including a set (as occurred on 2,251 of 2,664 rounds, 85%; dashed curve), and distribution for largest groups including a detected set (as occurred on 1,619 of the 2,251 rounds above, 72%; solid curve). (C) Mean fraction of rounds over all games in which subjects detected a set among the MAV cards, out of the rounds with such a set (solid curve), as compared with chance level (dashed curve), calculated (for each MAV-GS separately) by averaging the fraction of MAV sets from the total number of sets in the display, over all rounds. Actual findings are somewhat above the chance level, implying preference for detecting sets within the MAV group. (D) Mean response times (RTs) for detecting sets for four cases: when there was a set within the MAV and outside it (solid curves) and the detected one was within (black squares) or outside (gray diamonds); when there was a set only within the MAV (dashed black curve, squares) or only outside it (dashed gray curve, diamonds). Shorter RTs for sets within the MAV suggest that subjects show a preference for these more salient cards. Declining RTs with increases in MAV-GS suggest that MAV salience increases with group size.

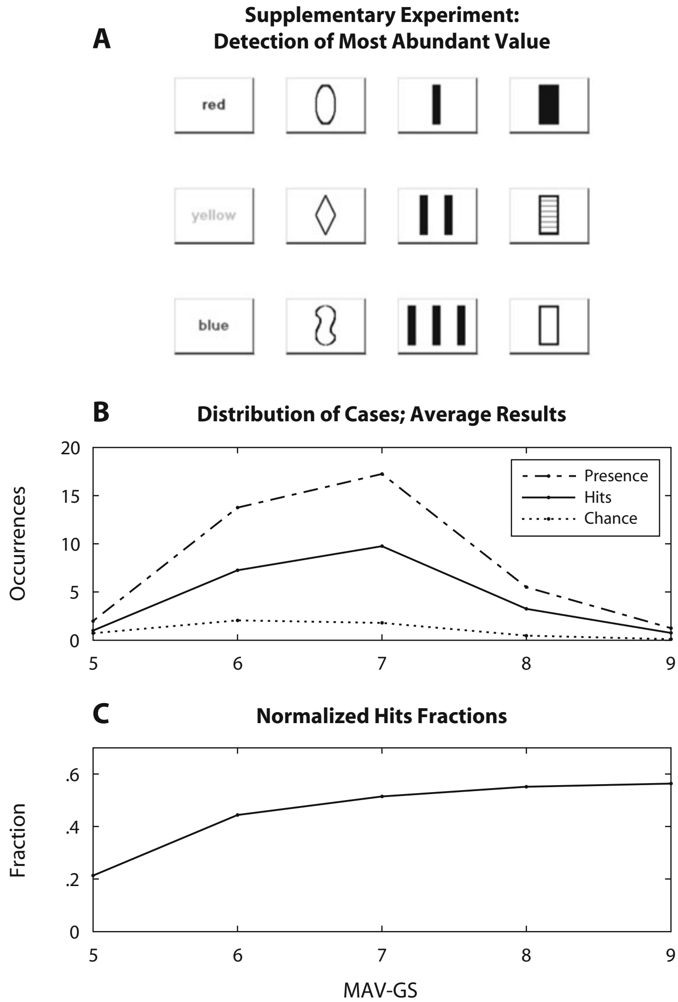

Figure 5. Supplementary task asking subjects to detect, "What is the most abundant value?" (A) Presentation of values for subject choice, following presentation of regular display of 12 cards (as in Figure 1). Each subject performed 40 rounds of this task. (B) Average results for 4 subjects as a function of most abundant value group size (MAV-GS): Number of correct answers (solid curve), as compared with number of occurrences (dashed curve) and with chance level (dotted curve), taking into consideration that several values can be MAVs in the same display. (C) Correct answers as fraction of distance from chance to number of occurrences of this MAV-GS, calculated by $(x - x_0)/(x_{max} - x_0)$, where x = hits, x_0 = chance, and x_{max} = presence. Thus, subjects had a good notion as to which was the MAV and could have used this information in playing the real game.

set within the MAV group (dashed curve). Experimental findings follow the chance-level increase with increased MAV-GS but are always somewhat above the chance level, suggesting that subjects indeed look preferentially within the MAV group, no matter what its size.

We also examined mean RT of detecting sets for different MAV-GSs, as plotted in Figure 4D. We examined four cases and compared them pairwise. When there is a set within the MAV group and another outside it, we can compare RTs when the detected set is entirely within the MAV group (solid black curve, squares) or outside it (solid gray curve, diamonds). Note that RTs are shorter for sets within the MAV group and that they decrease with increasing size, implying that the task becomes somewhat easier as the MAV-GS increases. We interpret these results as deriving from a preference in searching within the MAV group, so that sets there are detected more quickly. There are longer RTs for detecting sets outside the MAV group, not influenced by group size. Presumably, subjects waste time looking for a set within the MAV (so a set is detected outside the MAV only when it "can match the competition" of those inside). The MAV group itself may be more salient when it is larger. These suggestions are reinforced by comparing RTs when there is a set only within the MAV (dashed black curve, squares) or only outside it (dashed gray curve, diamonds). Again, for sets within the MAV group, RT decreases with increases in group size—that is, with MAV salience. As for sets outside the MAV group, RT increases with increases in group size (rather than being flat, like the solid gray curve). This increase may hint at search tactics, implying that the MAV cards distracted the search, by attracting attention to them, and only when a set was not found among them was a further search made.

We conclude that subjects preferentially search for sets within the MAV group, especially when the MAV-GS is large. Detecting a set within the MAV is faster than detecting one outside the MAV. Since a MAV group is a group of cards that are similar in a certain value, the larger the MAV-GS, the greater the degree of shared similarity, in the sense that there are more possible triplets with this dimensional similarity. This result therefore confirms the preference for similarity.

Perception of the MAV. In order to determine whether using the MAV is at all a feasible strategy for finding a set, we performed a supplementary experiment, testing the accuracy of detecting and reporting what is the MAV. This experiment was performed only after the subjects had played the regular game, in order not to influence the way in which they would play the game.

Twelve cards were displayed, as in the Set experiment, except that here, the display lasted only 5 sec. The task was to state which value was the most abundant. After the display disappeared, the 12 possible values were presented (as shown in Figure 5A), and the subject chose one, without time limitation. After the choice was made, another 12 cards were shown, with all the cards replaced (and not only 3, as in the regular game), so that the abundance of the values was entirely refreshed. Each subject performed 40 rounds of this task.

Average results for the 4 subjects are shown in Figure 5B. For each MAV-GS, we plot the average number of correct answers (solid curve), the average number of occurrences of this MAV-GS (dashed curve), and the chance level for correct answers (dotted curve), taking into account that several values could be the most abundant in the same display (thereby increasing the chance that one of them would be chosen). The average over subjects of correct answers is 22.25 (56%), nearly double the chance level of performance, which is ~12.5% (5/40) in total. Figure 5C shows correct answers as a fraction of the distance between the other two, by calculating $(x - x_0)/(x_{\max} - x_0)$, where x, x_0, and $x_{\max}$ are the measured correct answers, chance level, and total occurrences for each MAV-GS, respectively. There is a major increase from a MAV-GS of 5 to a MAV-GS of 6, and then it is pretty stable up to a MAV-GS of 9, with a value of ~0.5—that is, halfway between chance and the maximal possible number of correct answers.

Thus, the subjects had some notion of what was the most abundant value in the display, even though they were able to declare what it was in only a bit above half of the displays. Nevertheless, even though the MAV might not have been declared correctly (in the supplementary experiment), this still does not mean that this information was not used (even if unconsciously) for directing and speeding up set detection, as was found above. There is a difference between using information and being able to report it.

Place Effect

Do subjects find sets with cards close to each other more quickly and more often than they find sets with cards that are far from each other? As the total distance between the three cards of a set, we used the sum of the three Euclidian distances between each of the three pairs in the set (i.e., the triangle perimeter), as demonstrated in Figure 6A, in terms of the (vertical or horizontal) unit distance between adjacent display locations. There are 24 discrete distances.

Figure 6B shows the combinatorial (dashed line) versus the actual detection (solid line) probability of appearance of each distance. They show a very good fit. When they are plotted one versus the other (Figure 6C), the regression line has a slope of 1. We conclude that there was not much effect of the distance between cards (or the frequency of a particular distance). In addition, we find that there was no dependence of mean RT on distance (not shown). Taken together, these results suggest that subjects are able to perceive many cards at a single glance or that the order of scanning was hopping from place to place (perhaps on the basis of an attribute) and not necessarily to contiguous regions.

Influence of Location in the Display

Are there favored locations, which are more easily perceived by subjects? We analyzed the influence of a set including one (or both) of the two central locations. Figure 7A shows the frequency with which subjects actually chose a set with a card in each of the 12 locations (when there was more than one set in the display). As can be seen in the figure, the two central locations were slightly favored.

Figure 7B shows the frequency of choosing a set with a card as a function of the card's distance from the center of the display. There was a small, although nonsignificant, dependence, so that locations further from the center were favored less. We conclude that only the very central locations are somewhat favored.

Influence of Previous Set Card Locations

We wished to ensure that the locations of the cards included in the previously detected set did not affect triggering of the next set. More attention may have been paid to these locations because, here, the cards were replaced with new ones, while others remained from the preceding trial. On the other hand, other cards were already familiar, so that less attention may have been paid to the new cards. On the basis of 2,664 rounds, in .89 of the cases (2,372 rounds), there was a set that included one or more of these replaced card locations. This resembles the theoretical probability of .88, which we derive in Box 4.

Of these 2,372 rounds, in .506 of the cases, there existed at least one additional set, not including any previous location. We will consider only these 1,201 cases in which the subjects had a choice.

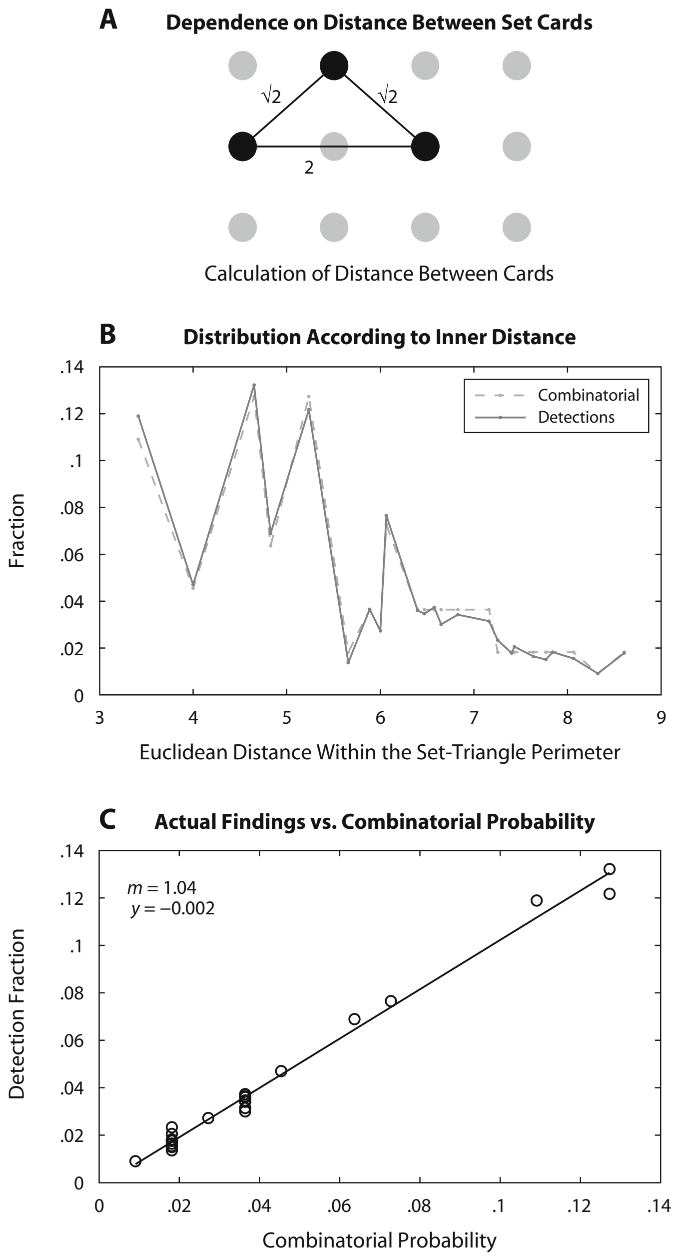

Figure 6. Effect of distance between cards. (A) Measurement technique. Calculation of distance between cards is done by counting as one unit the distance between two adjacent cards and summing all three distances, the triangle perimeter. There are 24 discrete distances, each appearing with a different probability. (B) Combinatorial distribution of distances (dashed lines) and actual distribution of detections (solid lines). Only (2,194) cases with at least two sets are considered. (C) Probabilities of detecting each distance versus combinatorial probabilities of presence of this distance. Note the good fit with a slope of 1, suggesting that there is no influence of distance among set cards on the probability of finding the set.

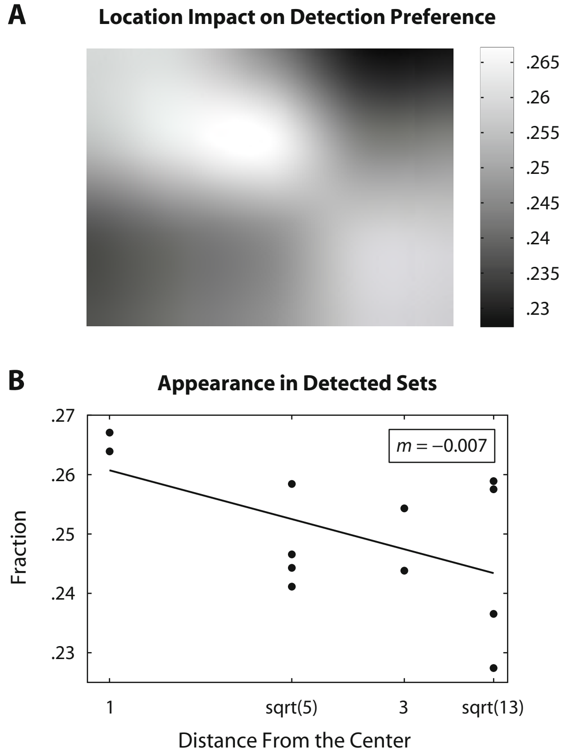

Figure 7. Influence of location in the display. (A) The fraction of appearances of each location in the detected sets, taking into consideration only the 2,313 rounds with more than one set present. Note that the fractions sum up to 3 (and not to 1), because of the three cards in a set. Note also that the value of .25 is the chance level. (B) Frequency of appearance of cards in the detected sets as a function of their distance from the center of the display, where distance is measured in units equal to half the vertical or horizontal distance between adjacent cards. $R^2 = .3$.

Averaging (over the 1,201 rounds) the empiric ratio of the number of sets including a previous location and the total number of sets present leads to .58 as the random choice level of choosing a set with a previous location. In practice, such sets were chosen in .67 of the cases (808 of 1,201 rounds).

In summary, in .89 of the rounds, there was a set including a previous location, similar to the theoretically expected value of .88. Looking only at rounds including both a set with a previous location and another set without such card, we found that in 67.28% (confidence interval, 67.0%–67.4%) of the cases, the chosen set was one including a previous location, as compared with a chance level of 58% (*SD*, ±14.16%; *SE*, ±0.41%). We conclude that there was some triggering by the newly placed cards (although not as much as there could have been—i.e., 100% of the 1,201 cases considered).

Conclusions for Experiment 1

Relating to similarity, we found that sets from lower classes were detected more quickly and more often (relative to their availability)—that is, with priority when more than one set was present simultaneously—suggesting that sets of lower classes are detected more easily and that greater perceptual similarity is a factor in set detection.

We found that the larger the number of sets present, the shorter the RTs. There was a good fit between the horse race model predictions and the actual results, suggesting that the model can account for our results. The implication of the success of the horse race model is independence (rather than synergy or interference) of the processes of finding each set when there is more than one present.

The MAV group may have been more salient when it was larger, reflected by decreasing RTs with increasing MAV-GS and by more distraction when there was no set there. There was some evidence of a preference in searching within the MAV group, supported by detections from the MAV group above chance level and by shorter RTs for sets within the MAV; presumably, the MAV cards distract the search, by attracting attention to them, so a set is detected outside of the MAV only when it "can match the competition" of those inside; or, in cases in which there is no set within the MAV cards, only when a set is not found among them is a further search made. This is the probable search strategy used. In detecting and reporting what was the MAV, the subjects' responses were halfway between chance and the maximal number of correct answers.

There was not much effect of the distance between the cards, suggesting that the subjects were able to perceive many cards at a single glance or that the order of scan was hopping from place to place, and not necessarily to contiguous regions. On the other hand, there was a slight preference for sets including (one of) the two central cards, those in the middle of the display.

There was some triggering by newly placed cards (previous set card locations); in a situation of choice, random behavior would predict a probability of 58% of choosing a set including a previous location, versus an actual occurrence of 67% of the cases. But this was still much less than it could have been (i.e., up to 100%).

EXPERIMENT 2
Dimensional Salience

If subjects find sets by first identifying similar cards (as is suggested by the results of Experiment 1), we might expect that sets with similarity in a more salient dimension (say, color) will be chosen over similarity in a less salient dimension (e.g., shape). We first ask whether there are more salient dimensions and then how their preference affects set identification. To this end, we performed a supplementary experiment to determine dimensional preference, in a subject-by-subject manner, expecting that the results may aid in understanding strategies used to detect sets, whether intentionally or not.

Method

We compare dimensional salience, using a graph theory algorithm to determine the ordering of the dimensions, and then examine how this ordering influences set detection. The salience of different dimensions may be compared in a straightforward manner by judg-

BOX 4
Combinatorics: Probability of at Least One of the Three Cards at the Replaced Locations Being Included in a Set

Instead of looking at the probability that these locations will be included in a set, we will look at the probability that the cards included in sets will include one of these cards, which is combinatorically the same. The latter is preferred because there are always exactly three previous locations, but if there is more than one set, the number of cards included in a set varies.

If there is one set in the 12-card array, the probability that *all* of the 3 cards in the set are *not* in *any* of the three locations of the newly placed cards is just

$$p = \frac{9}{12} \cdot \frac{8}{11} \cdot \frac{7}{10} = .38.$$

If there is more than one set in the array, involving x cards (five or more), the probability of the newly placed cards not overlapping any of these set-including cards is

$$p = \frac{12 - x}{12} \cdot \frac{11 - x}{11} \cdot \frac{10 - x}{10}.$$

Probability $(1 - p)$ by number of cards involved in sets (x) is shown in Table 1. Clearly, there is quite a high probability that one of the new cards will be included in an existing set. Of course, this is true for any group of three cards in the display, and players could as well concentrate on any convenient group, and not specifically on the replaced cards.

The empirical distribution P (based on the played games) of the number of cards in the display belonging to any present set is shown in the second row of the table.

Looking at the combinatorics and taking the dot product of the two rows of the table (the probability of there being x cards in the sets and their probability of including at least one of the three replaced cards) yields the weighted average probability that one of the present sets includes one of the three replaced card locations or, equivalently, that one of the three new cards is included in a set—namely, .88.

Table 1
Probability of at Least One Card of the Set Being New

	3	4	5	6	7	8	9	10	11	12
$(1 - p)$	.62	.75	.84	.91	.95	.98	.9955	1.0	1.0	1.0
P	.18	0	.17	.19	.13	.15	.10	.06	.02	.005

Note—$(1 - p)$, probability that at least one of x cards included in set(s) is among three changed cards. P, empirical probability that x cards in the display were involved in some set.

ing which of two test cards seems more similar to a reference card, implying that the dimensional change between the reference and the *other* card is more salient (Medin, 1973). The example of Figure 8A illustrates the method. We display three cards (each with only one element, and all of the same color): a filled oval reference card and empty oval and filled wave test cards, asking which seems more similar to the reference. If the filled wave is declared more similar, filling is more salient (since this is the changed dimension in the other stimulus), and vice versa.

The test was performed for all combinations of dimensions and values. Therefore, the total number of such comparisons is the product of the number of reference cards, v^d, the number of ways of choosing two dimensions, and the number of values in each of these two (leaving out the value of the reference card itself)—that is,

$$v^d \cdot \binom{d}{2} \cdot (v-1)^2 = 3^4 \cdot \binom{4}{2} \cdot (3-1)^2 = 1{,}944.$$

The outcome of these comparisons is translated into a fully directed graph with d nodes, representing the dimensions. Weights (w_{ij}) of the directed edges ($d_i \rightarrow d_j$), representing the salience, are assigned by the number of times dimension i is found to be salient over j, normalized to number of comparisons made between them. Then edges for which $w > .5$ are accepted, as demonstrated in Figure 8B.

We then require that the graph have a path through all nodes (a Hamiltonian path without closure). This condition ensures consistency. Because it is a full graph, this requirement is equivalent to requiring a directed acyclic graph (DAG). To fulfill this condition and find the ordering, the out-degree (d_{out}) of the nodes is sorted in descending order, from $(d - 1)$ to 0. When this forms a DAG, its path indicates the dimensional ordering. We wish to know whether this dimensional salience ordering relates to the detection of sets.

Results and Discussion

We tested 6 subjects following their playing several sessions of the usual Set game. The experimental results show within-subjects preference consistency but different orderings for different subjects. An example for 1 subject is shown in Figure 8B, and the average over all subjects in Figure 8C. Note that different people may regard different dimensions as salient, so before averaging, the dimensions were sorted according to preference for each subject, termed d1 . . . d4. Relative salience can also vary for the same person, without implying inconsistency, if two dimensions are compared on different *background* values of other dimensions. For example, if the reference shape is filled, the color may be more important than the shape, but if it is empty, it can be the reverse.

To measure the influence of dimensional salience on set detection, the dimensions were sorted according to pre-

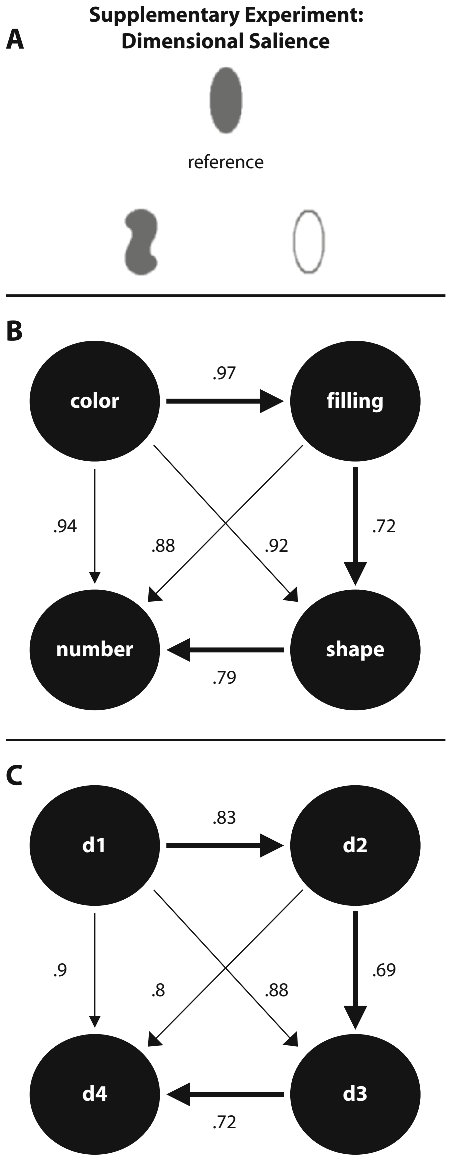

Figure 8. Dimensional salience and algorithm for determining order of individual dimensional preference. (A) The task. The subjects were shown a reference card and were asked to judge which of two test cards seemed more similar to it. Each test card differed from the reference on one dimension. If a certain card was chosen, this meant that the dimensional change between the reference and the *other* test card was more salient. In the example shown, if the left card seems more similar to the reference, filling is more salient than shape. (B) Demonstration of resulting directed acyclic graph (DAG) for 1 subject. Nodes represent dimensions, and directed edge weights represent fraction of times that one dimension was salient over the other. The path, indicating dimensional ordering, is created by sorting the out-degree (d_{out}) of the nodes in descending order. In this case, salience order was color, filling, shape, and number. (C) Average DAG over all the subjects. Since each subject had an individual preferred order, before averaging we sorted the dimensions according to each subject's preferred order, termed d1 . . . d4. Note that edges to less preferred dimensions have larger numbers (greater preferences).

ferred order for each subject separately, and the weights were averaged, in that order, over all subjects. This allowed analysis for all the subjects at once, even though each had a different ordering—for example, of the *i*th preferred dimension.

We then compared several parameters of detected and missed sets for each dimension, including the following: number of times there was similarity (or span) in each dimension, within the sets detected by the subject, in cases in which there was a choice among several sets (Figure 9A); for detected sets with similarity in each dimension, the number of sets present from lower classes (Figure 9B) or from the same class (Figure 9C), where we might expect dimensional salience to overcome preference for lower classes.

If dimensional salience significantly influences set detection, these parameters should vary systematically. However, we found no such monotonic dependence, so we infer that there is no apparent effect of dimension salience on set detection.

Relating the results of the two sections, so far, the three characteristics—class, MAV, and dimensional preference—are different and, as may be expected, play different roles in set detection. Class determines search procedure, and it turns out that finding sets of lower classes is easier than finding those of higher classes, perhaps because there is a natural preference for perceiving similarity. MAV is a characteristic of the cards in the display and, in accord with the preference for similarity, plays a role in the finding of sets (in the MAV group, thus sharing similarity in its value). In contrast to these characteristics, dimensional preference is a personal preference (which we found varies from subject to subject) and has no bearing, on average, on set detection success. In addition, there is usually a conflict between detection based on MAV and that based on dimensional preference, and MAV wins. Thus, it may not be surprising that personal preferences do not determine performance in the long run.

This division between different levels of influence of different perceptual aspects of the stimulus array may reflect the special status of the Set game: As was described in the introduction, Set is very basically a perceptual game, in that it depends on perceiving combinations of values among the 12 presented cards. On the other hand, the task of the game is conceptual, in that specific cognitive rules must be followed. Nevertheless, as has been mentioned, conceptual processes may also, in turn, derive from and be influenced by perceptual attributes (Goldstone & Barsalou, 1998). As such, it is natural that different perceptual aspects will have different levels of influence on the processes underlying performance of this complex task.

EXPERIMENT 3
Learning and Generalization

In Experiment 3, we tested the dynamics of learning the Set game. We also asked whether training-induced learn-

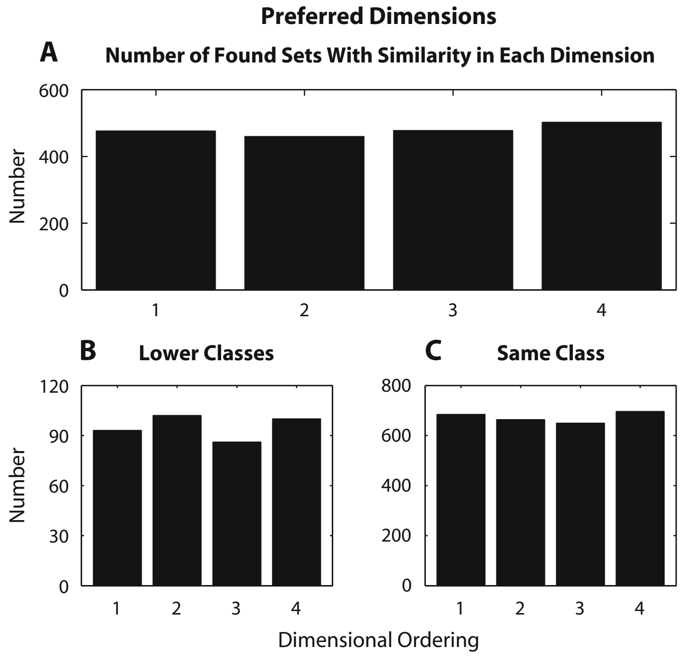

Figure 9. Influence of dimensional preference. Dimensions were sorted according to individual preference, as in Figures 8B and 8C. Only cases in which there was a choice among several sets were considered here. (A) Total number of times for all the subjects that the set detected had similarity on each of the four dimensions, ordered according to the preference of each subject. (B and C) Number of sets present from *lower* classes (B) or the *same* class (C), when the detected set had similarity in each dimension ordered by individual preference.

ing would generalize to playing the game with changed stimulus values (see, e.g., Figure 1C). That is, after training improved performance, would changing stimulus values start learning all over again, or would the ability to identify a set be now established for all stimuli? We were interested in whether learning this task is high or low level. If there is generalization and transfer of learning effects when playing with new stimuli, the implication is that learning is a high-level effect, whereas if training is specific to trained stimuli, learning may be a low-level effect (Ahissar & Hochstein, 1997, 2004; Hochstein & Ahissar, 2002).

Method

Learning experiments included three sessions with 9–12 games. Ten subjects participated in three complete sessions, with 1 or 2 games in the first session, 3–5 in the second, and 3–6 in the third. Following three sessions with the original cards, transfer was tested for 6 of the subjects for cards with shapes changed to a circle, a triangle, and a square and with changed colors, as demonstrated in Figure 1C.

Results and Discussion

The example learning curve in Figure 10 shows mean RT for each class as the games proceed. There is a gradual improvement, seen as a decreasing RT. Again, there are class-dependent characteristics (see above), with more stabilization and lower RTs for lower classes. The arrow points to the time when the different version was applied. There is not much difference in the RTs after this point.

We tested the average across all 6 subjects who played the generalized game, taking their first three games, the last three before generalization, and those just after the new version was applied (as shown in Figure 11). There was highly significant learning from the first to the third game (one-tailed paired t test over subjects, between the two games, with $p < .005$); also highly significant is the difference between performance on the first game and that on the third-to-last game played before the subjects switched tasks—that is, the (L-2) game (one-tailed paired t test, $p < .001$). There was no increase in RT when the subjects moved to the new version (one-tailed paired t test yields $p = .3$).

We conclude that the training effect generalized to playing the game with new values. This training generalization may have resulted from the fact that the subjects did not yet reach a stabilized “automatic” level (Goldstone, 1998; Treisman, Vieira, & Hayes, 1992)—an assumption supported by the mean RT—and that as long as performance of this task depended on a cognitive process, training generalized to different values.

SUMMARY AND GENERAL DISCUSSION

Summary

We found that when subjects played the Set game, several parameters influenced set detection (Experiment 1), including the following: similarity in values (within a

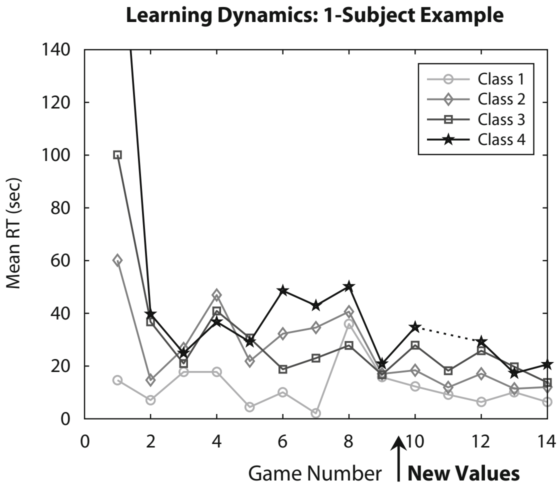

Figure 10. Learning and generalization: Example data for 1 subject. Learning dynamics are shown by mean response time (RT) for each class as the games proceed. The arrow points to the place where the different version (Figure 1C) was used.

dimension); number of existing sets in the display, with RT acting according to the horse race model, implying independence of simultaneous searches; and the MAV and its group size, which was searched preferentially, also confirming the preference for similarity.

We used an algorithm for determining dimensional salience (Experiment 2) on the basis of direct comparisons, using graph theory. The subjects showed a consistent but individual order of preference for dimensions, but this seems not to have affected set identification preference.

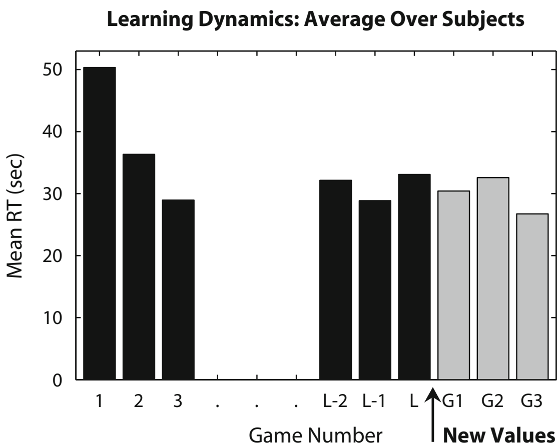

Figure 11. Learning and generalization: Average results for 6 subjects. Mean response times (RTs) for the first and last three games with the original version (regardless of number played), and for the first three games with a different version (gray bars). Note the significant learning from the first to the second game and the lack of increase in RT when the subjects moved to a new version. The *SD* for RT over games changes in the same manner (not shown).

There was a gradual improvement in speed of playing the Set game with experience, with class-dependent characteristics (Experiment 3). Training-induced learning generalized to new versions of the game with new stimuli, suggesting a high-level learning effect.

These results were enabled by "complications" inherent in the Set game but not present in other categorization tasks (see the introduction), such as the possibility for having more than one set present in each display (allowing the study of *competition* among the processes seeking them) and the game's including a *span* rule for dimensions on which similarity is not found. This forced the subjects to intermix conceptual and perceptual aspects in the search for the rules applying to the current display, or even to each of the sets present in it. It also allowed us to find perceptual influences in this conceptual task. These implications will be discussed below.

An interesting question, regarding perception in general, is what happens in the brain during the very recognition of a set, at the exact moment of conscious perception, often called the moment of insight (Ahissar & Hochstein, 1997; Bowden & Jung-Beeman, 2003; Rubin, Nakayama, & Shapley, 1997; Smith, Gosselin, & Schyns, 2006). Future studies using the Set game as an interface may address this issue.

General Discussion

The Set game task is complex because it involves both perceptual and cognitive features (cf. Pomerantz, 2002; Schyns, Bonnar, & Gosselin, 2002). Subjects must perceive the values and the relationships among the cards on the basis of four visual dimensions, but they must decide which three of the present cards form a set on the basis of cognitive rules. Improvement may come from improved or faster perception of the dimensional values present, from better understanding and application of the cognitive rules, or from a combination of these.

Even though the Set game is not a game of categorization, it may be of value to compare these two tasks. Regarding our finding that sets with more similarity were found more often and more rapidly than others, we note that categorization, too, may depend on finding similarities among different elements (Goldstone, 1994). For example, categorization has been seen as the finding of a prototype or group of exemplars and the other objects that are more similar to these than to its competitors (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). This includes similarities along a number of dimensions (referred to as *family resemblance* when there is similarity along many but not all dimensions; Medin, Wattenmaker, & Hampson, 1987; Regehr & Brooks, 1993; Rosch & Mervis, 1975), although subjects may base their categorization on a single salient dimension (Ashby, Queller, & Berretty, 1999; Bower & Trabasso, 1963). It is possible that the same basic mechanisms that underlie categorization are also used in playing the Set game. If this is the case, the bias that we found toward lower class sets may reflect the tendency for real-world categories to be organized around similarities, rather than around differences (Ashby & Maddox, 2005; Hammer et al., 2005, 2007, in press; Medin et al., 1987;

Rosch et al., 1976). Indeed, Hammer and colleagues found a preference for using common feature regularities, rather than distinctive feature irregularities, supporting the conclusion that similarity is more beneficial and, ultimately, more natural than is dissimilarity for use in categorization. Similarity perception also develops earlier than learning from differences (Hammer et al., 2008). A between-object similarity-detecting mechanism developed for categorization may serve set detection as well. Such a mechanism may involve, for example, mutual enhancement between local feature detectors. (We built a neural network model containing a mechanism working on this principle, and it, too, found mainly low-class sets.) We conclude that the preference for finding sets with more similarity provides supporting evidence for the presence of a similarity-detecting mechanism.

Returning to the issue of the perceptual and cognitive aspects of playing the Set game, we now ask what role perception plays in this cognitive task. Is the task purely conceptual, or does it contain a perceptual element? There are at least two factors that suggest that perception does play an important role in set recognition, as follows. We found that seemingly inconsequential perceptual features influenced which sets were detected (when more than one set was present) and the speed of detecting a set (even when there was only one). For example, there was an important dependence on the MAV among the 12 possible values (3 per dimension for four dimensions) in the display. The second factor that indicates a perceptual element in the cognitive task of set detection relates to the issue of similarity versus span. Finding a set depends on detecting a mixture of both the more perceptual similarity and the more conceptual span. Although the rule of similarity can be seen as just as cognitive a rule as the rule underlying a span, nevertheless, only similarity detection can be seen also as a basic, immediate, and automatic perceptual mechanism. This may be the source of the bias that we found for detecting sets with more similarity, rather than with more span. (In fact there would be an inherent advantage to detecting spans, since on average, there is much more spanning than similarity in the possible sets.) Thus, the present results confirm that even in a task such as Set, which inherently depends on both similarity and span detection—and thus, on conceptual processes—the overwhelming influence of the perceptual determines that priority will be given to sets with greater similarity.

The experimental result that subjects find sets of lower classes preferentially and more quickly suggests that people may have built-in mechanisms for finding similarities, but not for finding spans—that is, groups of items that do not include two that are similar along the relevant dimension. In this case, finding a span may be a cognitive, analytic, or abstract reasoning task. Our result that even in this case, subjects are sensitive to perceptual features of an array (such as the MAV) suggests that perceptual and cognitive functions may not be totally separate (see Goldstone & Barsalou, 1998; Landy & Goldstone, 2007).

Overall, one of our major findings is that people perceive similarity within a dimension better than difference—that is, a span. This might seem, at first, to contradict the fact that the visual system detects change. The brain specializes in identifying difference. This is why we are so good at detecting an object that differs from its surround. Why, then, are we so good at identifying three similar objects among a mixture of items? This could be because, among such a diverse collection, what is unique is similarity. When the environment is unified, we detect difference; when the environment is diverse and colorful, we detect the few points of similarity within it.

AUTHOR NOTE

This research was supported by grants from the U.S.–Israel Binational Science Foundation (BSF) and the Israel Science Foundation (ISF). We thank Anne Treisman for inspiring discussions during the course of this study. We thank Daphna Weinshall and Leon Deouell for constructive support and Ron Katz for helpful conversations and review. We thank reviewer James Pomerantz for pointing out the double meaning of the term *similarity*. We thank him and other (anonymous) reviewers for useful comments.
Correspondence concerning this article should be addressed to M. Jacob, Department of Neurobiology, Institute of Life Sciences, Hebrew University, Jerusalem 91904, Israel (e-mail: michaljb@gmail.com).

REFERENCES

AHISSAR, M., & HOCHSTEIN, S. (1997). Task difficulty and the specificity of perceptual learning. *Nature*, **387**, 401-406.

AHISSAR, M., & HOCHSTEIN, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, **8**, 457-464.

ASHBY, G. F., & MADDOX, T. W. (2005). Human category learning. *Annual Review of Psychology*, **56**, 149-178.

ASHBY, G. F., QUELLER, S., & BERRETTY, P. M. (1999). On the dominance of unidimensional rules in unsupervised categorization. *Perception & Psychophysics*, **61**, 1178-1199.

BERG, E. A. (1948). A simple objective technique for measuring flexibility in thinking. *Journal of General Psychology*, **39**, 15-22.

BOWDEN, E. M., & JUNG-BEEMAN, M. (2003). Aha! Insight experience correlates with solution activation in the right hemisphere. *Psychonomic Bulletin & Review*, **10**, 730-737.

BOWER, G., & TRABASSO, T. (1963). Reversals prior to solution in concept identification. *Journal of Experimental Psychology*, **66**, 409-418.

CARR, T. H. (1992). Automaticity and cognitive anatomy: Is word recognition "automatic"? *American Journal of Psychology*, **105**, 201-237.

CORBALLIS, M. C. (1998). Interhemispheric neural summation in the absence of the corpus callosum. *Brain*, **121**, 1795-1807.

EGETH, H. E., & MORDKOFF, J. T. (1991). Redundancy gain revisited: Evidence for parallel processing of separable dimensions. In G. R. Lockhead & J. R. Pomerantz (Eds.), *The perception of structure: Essays in honor of Wendell R. Garner* (pp. 131-143). Washington, DC: American Psychological Association.

GARNER, W. R., & LEE, W. (1962). An analysis of redundancy in perceptual discrimination. *Perceptual & Motor Skills*, **15**, 367-388.

GOLDSTONE, R. L. (1994). The role of similarity in categorization: Providing a groundwork. *Cognition*, **52**, 125-157.

GOLDSTONE, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, **49**, 585-612.

GOLDSTONE, R. L., & BARSALOU, L. W. (1998). Reuniting perception and conception. *Cognition*, **65**, 231-262.

GRAHAM, N. V. S. (1989). *Visual pattern analyzers*. New York: Oxford University Press.

GRANT, D. A., & BERG, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card sorting problem. *Journal of Experimental Psychology*, **38**, 404-411.

HAMMER, R., DIESENDRUCK, G., WEINSHALL, D., & HOCHSTEIN, S. (2008). *The development of category learning strategies: What makes the difference?* Manuscript submitted for publication.

HAMMER, R., HERTZ, T., HOCHSTEIN, S., & WEINSHALL, D. (2005). Category learning from equivalence constraints. *Cognitive Sciences*, **27**(Suppl.), 893-898.

Hammer, R., Hertz, T., Hochstein, S., & Weinshall, D. (2007). Classification with positive and negative equivalence constraints: Theory, computation and human experiments. In F. Mele, G. Ramella, S. Santillo, & F. Ventriglia (Eds.), *Brain, vision, and artificial intelligence* (pp. 264-276). Berlin: Springer.

Hammer, R., Hertz, T., Hochstein, S., & Weinshall, D. (in press). Category learning from equivalence constraints. *Cognitive Processing*.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. C., & Curtiss, G. (1993). *Wisconsin Card Sorting Test manual, revised and expanded*. Odessa, FL: Psychological Assessment Resources.

Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron*, **36**, 791-804.

Holyoak, K. J. (1990). Problem solving. In D. N. Osherson & E. E. Smith (Eds.), *Thinking: An invitation to cognitive science* (Vol. 3, pp. 117-146). Cambridge, MA: MIT Press.

Koffka, K. (1935). *Principles of Gestalt psychology*. New York: Harcourt Brace.

Köhler, W. (1929). *Gestalt psychology*. New York: Liversight.

Landy, D., & Goldstone, R. L. (2007). How abstract is symbolic thought? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **33**, 720-733.

Medin, D. L. (1973). Measuring and training dimensional preferences. *Child Development*, **44**, 359-362.

Medin, D. L., Wattenmaker, W. D., & Hampson, S. E. (1987). Family resemblance, conceptual cohesiveness, and category construction. *Cognitive Psychology*, **19**, 242-279.

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, **14**, 247-279.

Miller, J. (1986). Timecourse of coactivation in bimodal divided attention. *Perception & Psychophysics*, **40**, 331-343.

Monnier, P. (2006). Detection of multidimensional targets in visual search. *Vision Research*, **46**, 4083-4090.

Oliva, A., & Torralba, A. (2006). Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research*, **155**, 23-36.

Pomerantz, J. R. (2002). Perception: Overview. In R. A. Wilson & F. C. Keil (Eds.), *Encyclopedia of cognitive science* (pp. 527-537). London: Nature Publishing.

Raab, D. H. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*, **24**, 574-590.

Regehr, G., & Brooks, L. R. (1993). Perceptual manifestations of an analytic structure: The priority of holistic individuation. *Journal of Experimental Psychology: General*, **122**, 92-114.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, **7**, 573-605.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, **8**, 382-439.

Rubin, N., Nakayama, K., & Shapley, R. (1997). Abrupt learning and retinal size specificity in illusory contour perception. *Current Biology*, **7**, 461-467.

Schyns, P. G., Bonnar, L., & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science*, **13**, 402-409.

Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time and spatial scale dependent scene recognition. *Psychological Science*, **5**, 195-200.

Smith, M. L., Gosselin, F., & Schyns, P. G. (2006). Perceptual moments of conscious visual experience inferred from oscillatory brain activity. *Proceedings of the National Academy of Sciences*, **103**, 5626-5631.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, **18**, 643-662.

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge: Cambridge University Press.

Treisman, A., Vieira, A., & Hayes, A. (1992). Automaticity and preattentive processing. *American Journal of Psychology*, **105**, 341-362.

Tversky, A. (1977). Features of similarity. *Psychological Review*, **84**, 327-352.

Ullman, S. (1984). Visual routines. *Cognition*, **18**, 97-159.

(Manuscript received April 10, 2007;
revision accepted for publication March 17, 2008.)