Editorial Manager(tm) for Cognitive Science
Manuscript Draft

Manuscript Number:

Title: Set as an Instance of a Real-world Visual-cognitive Task

Article Type: Regular Article (12,000 words)

Keywords: cognitive architecture; visual attention; cognitive control; games; ACT-R; problem solving.

Corresponding Author: Enkhbold Nyamsuren, Ph.D. candidate

Corresponding Author's Institution: University of Groningen

First Author: Enkhbold Nyamsuren, Ph.D. candidate

Order of Authors: Enkhbold Nyamsuren, Ph.D. candidate;Niels A. Taatgen, Professor

Abstract: The paper investigates how people accomplish tasks requiring complex visuospatial cognition and conceptual reasoning in the game SET. Our eye-tracking study with SET shows that it is not a purely perceptual game, but a game with a complex top-down planning influenced by bottom-up perceptual features. Combinatorics and statistical regression analysis of eye-movement protocols show how the different perceptual aspects of the game influence the strategy. Furthermore, they show a gradual shift during a game from bottom-up similarity-based processes to top-down conceptual processes. The assumptions are validated with an ACT-R model of a human SET player implemented on the basis of experimental findings. Finally, we discuss how the findings in this paper can be generalized to the real-world problem solving tasks where the bottom-up reactive behavior to the events in the external world and visuospatial stimuli is as equally important as top-down planning.

Set as an Instance of a Real-world Visual-cognitive Task

Enkhbold Nyamsuren and Niels A. Taatgen

University of Groningen

Author Note

Enkhbold Nyamsuren, Department of Artificial Intelligence, University of Groningen; Niels A. Taatgen, Department of Artificial Intelligence, University of Groningen.

Correspondence concerning this article should be addressed to Enkhbold Nyamsuren or Niels A. Taatgen, Department of Artificial Intelligence, University of Groningen, Nijenborgh 9, 9747 AG Groningen, Netherlands.

E-mail: e.nyamsuren@rug.nl; n.a.taatgen@rug.nl

Abstract

The paper investigates how people accomplish tasks requiring complex visuospatial cognition and conceptual reasoning in the game SET. Our eye-tracking study with SET shows that it is not a purely perceptual game, but a game with a complex top-down planning influenced by bottom-up perceptual features. Combinatorics and statistical regression analysis of eye-movement protocols show how the different perceptual aspects of the game influence the strategy. Furthermore, they show a gradual shift during a game from bottom-up similarity-based processes to top-down conceptual processes. The assumptions are validated with an ACT-R model of a human SET player implemented on the basis of experimental findings. Finally, we discuss how the findings in this paper can be generalized to the real-world problem solving tasks where the bottom-up reactive behavior to the events in the external world and visuospatial stimuli is as equally important as top-down planning.

Keywords: cognitive architecture, visual attention, cognitive control, games, ACT-R, problem solving

Human performance in complex tasks is often a combination of internal planning and responding appropriately to the environment. Nevertheless, cognitive models of complex tasks typically focus on the mental planning aspects, and fail to take into account that the external world can heavily influence the control of behavior.

The role of the environment was first recognized in robotics (Brooks, 1991) but was later extended to human cognition in the *embodied cognition* approach (e.g., Clark, 1997; Kirsh & Maglio, 1994). The challenge is to understand how control is shared between goal-driven planning and processes that are driven by perceptual input. The approach we will take in this paper is to assume two parallel processes: a bottom-up visual process that scans the visual field on the basis of saliency and similarity, and a top-down planning process that tries to achieve the goal, but also biases bottom-up process.

Finding an appropriate task to study the cognitive aspects of human behavior in real-life situation is not easy. However, games provide environments that often require the same type of complex processes that are usually involved in real-world situations (Green & Bavelier, 2004). This has the advantage that the behavior of a player can be studied in a controlled environment. These qualities make games on a computer an ideal tool for studying complex cognitive processes. One such game is the card game SET[1].

The SET card deck consists of 81 cards. Each card differs from other cards by a unique combination of four attributes: color, number, shape and shading. Each attribute can have one of three distinct values: red, green and blue for the color attribute; open, solid and textured for the shading attribute; one, two and three for the number attribute; oval, rectangle and wiggle for the shape attribute. The gameplay for SET is relatively simple. At any moment in the game, 12 cards are dealt open, as it is shown in Figure 1. From those 12 cards, players should find any combination of three cards, further referred to as a set, satisfying the main rule. This rule states that in the three cards the values for a particular attribute should be all the same or all different. The number of different attributes in the set cards is further referred as the *level* of the set. As such, the set, in which only one attribute is different, is a level 1 set. Correspondingly, there can be sets of level 2, 3 or 4. Figure 1 shows examples of level 1 (different shape) and level 4 sets (all attributes are different). In the regular game, if a player finds a set, he or she picks up the three cards that form the set, and replaces them with new cards from the deck. After the deck runs out the player with most cards wins. Even though a regular game of set consists of multiple rounds, we will refer to a “game of set” in what is normally a single round: finding a set in 12 displayed cards.

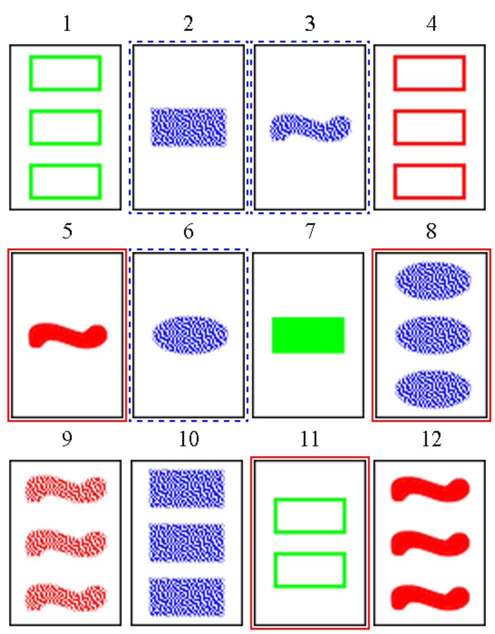

Figure 1. An example array of 12 cards. The cards with the solid highlight form level 4 set (all attributes are different), and cards with dashed highlight form level 1 set (Shape is different, and all other attributes are the same).

There are several advantages of choosing SET as a target game of study. One of the appealing parts of SET comes from the simplicity of the game dynamics. The game has very simple rules to follow and a relatively static game environment. Despite the simplicity, SET requires complex cognitive processes including pattern recognition, visuospatial processing and decision making. It is our hypothesis that in SET both cognitive and perceptual processes are equally important to play the game. As such, the game of SET provides an excellent opportunity to study the dynamics of such processes in a relatively simple game environment.

Finally, the game is unpredictable requiring dynamic and real-time decision making. There are $7*10^{13}$ possible combinations of 12 cards and 220 possible choices of three cards out of the array of 12 cards. It makes a detailed strategy planning impossible. In some regards SET is quite similar to Tetris (Kirsh & Maglio, 1994). In Tetris the player's behavior is not purely determined by strategy, because the player has to react to the next available block. Similarly, in SET players cannot really decide the strategy unless all the cards are observed. Players have to come up with the strategy on the fly after viewing the cards. Furthermore, the need to find the sets of different levels forces players to change the strategy as the game progresses. Such dynamic and unpredictable nature of the game makes SET an interesting target of a study.

Related Work

There are two other studies directly relevant to the work in this paper. Jacob and Hochstein (2008) did several experiments with human subjects playing SET on a computer without any opponent. Each experiment was designed to test a particular aspect of the game including a strategy of playing the game, dependency of the performance on the set level, attribute preference and the learning. Taatgen, Oploo, Braaksma and Niemantsverdriet (2003) also did similar experiment aimed for studying the strategy of playing the game and developed a computer model of a human player.

Strategy of playing SET

Jacob and Hochstein (2008) demonstrated that SET players prefer to look at perceptually similar cards, and, for the comparison of the cards, mainly rely on the perceptual processes such as similarity-detecting process. According to the authors, bias to the perceptual similarity and corresponding bottom-up processes can explain why players need less time to find lower level sets than higher level sets. Taatgen et al. (2003) also reached the conclusion that the perceptual elements play a greater role in finding lower level sets. They suggested a strategy where a player looks at an arbitrary first card then at a second card that shares an attribute value. Next, the player predicts the third card and determines whether that card is one of the remaining ten cards. Taatgen et al. (2003) also hypothesized that the choice of the first card might not be arbitrary in some cases. They proposed that players try to find the set among the cards that have attribute value occurring in more than half of 12 cards. For example, if there are many red cards, it is attractive to search for a set among those cards. Taatgen et al. (2003) implemented this strategy in an ACT-R model. However, the data they collected did not have enough detail to determine whether subjects use such a strategy.

Jacob and Hochstein (2008) proposed a generalization of the above strategy based on the notions of the most abundant value and the most abundant value group. The former refers to an attribute value that occurs most, and the latter refers to the group of cards that have the most abundant value. They found that the sets belonging to the most abundant value group are preferred to the sets outside of that group. In addition, the time required to find the set in the most abundant value group decreased as the size of the group increased. Most abundant value group was preferred to any other value group independently of the attribute type. Jacob and Hochstein (2008) suggested a *dimension reduction strategy* where players try to reduce the four dimensional search space into three by choosing to look at cards that have one or more attribute values in common. It was further claimed that this dimension reduction strategy is primarily used with the most abundant value.

Learning in SET

There is very little discussion on the aspects that result in a difference between novice and expert players. Jacob and Hochstein (2008) briefly investigated the learning in SET. They did a separate experiment to examine whether the training-induced learning can be generalized to find the sets that were not presented during the training. Indeed, the analysis of the subjects' reaction times indicated that such generalized learning occurs. However, very little discussion was provided on the nature of the learning and on the processes involved in it. On the other hand, Taatgen et al. (2003) argued that expert players have optimized comparison process of cards. Such optimization happens through the gradual transition from the declarative knowledge to the procedural knowledge resulting in a faster comparison of the cards. The corresponding model was able to learn through proceduralization and make a transition from the novice player to the expert player. Although the model is simple and relies mainly on higher-level cognitive processes, it was able to reproduce human players' reaction times with a reasonable degree of approximation. In addition, the model was able to generalize the knowledge obtained through the learning to find the sets it had never seen before. Such learning is consistent with the Jacob and Hochstein's findings about generalizable learning.

Research Objectives

Taatgen et al. (2003) used questionnaires and reaction times to gain understanding about player's behavior. Jacob and Hochstein (2008) used combinatorial analysis of reaction time. Although both studies suggest players use a dimension reduction strategy on the most abundant attribute value, this cannot be the complete explanation because a set cannot always be found in that group. In particular, level 4 sets have no attributes at all in common, making them impossible to find with that strategy. To gather real-time behavioral data that can provide more insight into previously hidden aspects of user behavior, we decided to use eye tracking. Since many studies have shown that the eye movement protocols directly or, at least, indirectly reflect both the cognitive processes and the amount of cognitive load (Kong, Schunn, & Wallstrom, 2010; Rayner, 1995; Salvucci, 1999), we considered the eye tracking a viable choice for studying human behavior.

Cognitive and Perceptual Processes

It was discussed earlier that SET players heavily rely on bottom-up perceptual processes. However, despite the importance of the perceptual elements, it is obvious that players have to utilize top-down processes to find the higher level sets. Players cannot rely anymore on perceptual similarity and have to deliberately search for the dissimilar cards. When and how the two types of processes are used is still an open question. In fact, as we suggest, this question is especially relevant within the context of player's experience.

In the game of SET, there is a significant gap between performance of novice and expert players. We would like to investigate what aspects at the cognitive and the perceptual elements result in the differences between the two groups of players.

Taatgen et al. (2003) argued that novice and expert players differ little in performance when it comes to finding lower level sets. However, they differ significantly in reaction times for finding higher level sets. One explanation for this effect might be that all players are likely to rely on perceptual processes to find lower level sets (Jacob & Hochstein, 2008). However, at same time players require explicit cognitive control to find higher level sets, and it is our assumption that more experienced players are better at top-down control.

It is possible that novice players rely more on perceptual processes for decision-making, while expert players rely more on conceptual aspects of the game. For example, for novice players, the choice of the cards to look at might be driven by the perceptual similarity, in contrast. The expert player, on the other hand, might be driven by both the perceptual similarity and top-down process, such as specific strategy.

Correspondingly, it is interesting to study the type of learning processes that lead to the emergence of the difference between expert and novice players. Within the context of the previous assumption, we expect to some kind of learning processes that leads to shift of balance from bottom-up processes to top-down processes. In other words, we expect a learning process where the involvement of top-down processes increases gradually, while the role of bottom-up processes decreases also gradually. We are planning to investigate whether such transition exists and how it occurs.

Improved ACT-R Model

The ACT-R model created by Taatgen et al. (2003) was able to closely approximate the human player's reaction times. Its main drawback is that it fully predicts the third card given the first two cards it has looked at, and then searches for that card among the remaining cards. It therefore does fully use a dimension reduction strategy, and also does not use perceptual similarity to find sets. In other words, it uses a pure top-down strategy. Our aim is to test whether a model with greater emphasis on perceptual elements of the game can explain the human data.

Experiment

Subjects

In total, 14 subjects participated in the experiment. The age of the subjects ranged from 20 to 30 years. All subjects were either students or staff members of University of Groningen. The subjects' previous experience with SET varied greatly: from few played games to several years of experience. Hence, the reaction times were chosen as an indicator of subject's overall experience.

Design and Procedure

Every subject was asked to do 60 trials. The group of 60 trials was same for all subjects, but the order of trials was determined randomly for every subject. Each trial consisted of 12 cards shown on a computer screen and arranged to an array similar to one shown in Figure 1. Each trial had exactly one combination of three cards that formed a set. Subjects were aware that there is always only one set in each round, but where not told about the level of the set.

All 60 trials were randomly generated with the constraint that all four levels of difficulty were equally represented in the experiment. In 30 trials one of the set cards was highlighted with a red border. These trials were distributed evenly over the four levels, with 7 or 8 trials of each level for each of the two highlighting conditions. The highlighted card belonged to the set and served as a clue for the subject to find the other two cards. Subjects were explained about the meaning of the highlighted card during the practice session. The presence of the highlighted card should make the task of finding the set much easier. For example, it decreases the number of possible unique combinations of three cards from 220 to 55. This is a four times reduction in complexity of the problem in terms of the search space.

In each trial, the subject was asked to find the cards belonging to the set and select them with the mouse. After successful selection of all three set cards or expiration of a time limit of 180 s, the next trial was automatically shown to the subject. In case of failure to find the set, the reaction time for that trial was recorded as 180 s. The sequence of trials was determined randomly for every subject.

Before starting the experiment, the subjects were asked to do four practice trials, one from each level. The results from the practice trials were not included in the analysis.

Eye Tracking

An EyeLink 1000 eye tracker was used for recording the eye movements. It is a desktop-mounted remote eye tracker with monocular sampling rate of 500Hz and spatial resolution of < 0.01° RMS. The card images were shown on 20 inch LCD monitor with screen size of 1024×768 pixels and screen resolution of 64 pixels/inch. The card images had size of 124×184 pixels. The horizontal and vertical distances between images were 80 and 70 pixels respectively. The approximate viewing distance is 70 centimeters. The gaze position was calculated using eye's corneal reflection captured with infrared camera and compensated for the head movement. The eye tracker's default parameters were used to convert gaze positions into events such as fixations and saccades. The calibration of an eye tracker was performed at the start and during the experiment, if necessary, with the accuracy of 0.8° being considered as an acceptable measure. After each trial the subject was asked to do drift correction as an additional corrective measure.

Experiment Results

Reaction Times

As it is shown in Figure 2, subjects differed significantly in terms of the mean reaction times, reflecting their different levels of expertise in SET. As can be seen from the graph, there are roughly three groups of subjects: the group of intermediate players that form a plateau at the middle of the graph, and groups of expert and novice players that are at the left and right hand sides of the plateau respectively.

Figure 3 shows the reaction times by level and highlighted condition. It shows that having a highlighted card as a clue decreases the reaction time by more than twice. This effect can be observed in all three groups of subjects and in all levels. Secondly, it is very clear that reaction time is largely dependent on the level. The very same phenomenon was also observed in previous studies (Taatgen et al., 2003; Jacob & Hochstein, 2008). This effect suggests that there is some correlation between visual similarity of the cards in a set and the time required to find the set. Subjects require less time to find sets in which cards are visually similar. This suggests that participants apply a similarity-based strategy like the dimension reduction. However, it is our hypothesis that dimension reduction is much more common strategy and not restricted to the most abundant value. In the following subsection we have investigated for evidence that the dimension reduction strategy was applied by the subjects.

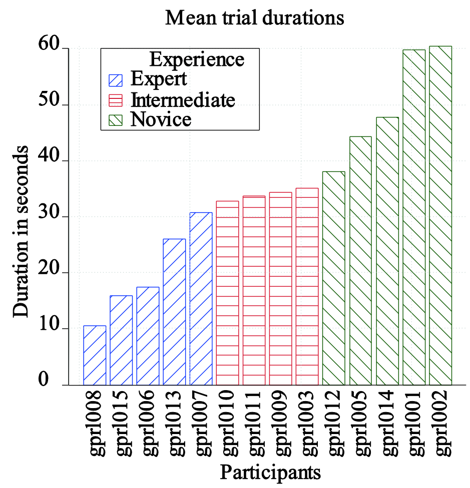

Figure 2. The mean reaction times averaged over all trials for each subject. The subjects are divided into three groups of expert, intermediate and novice players.

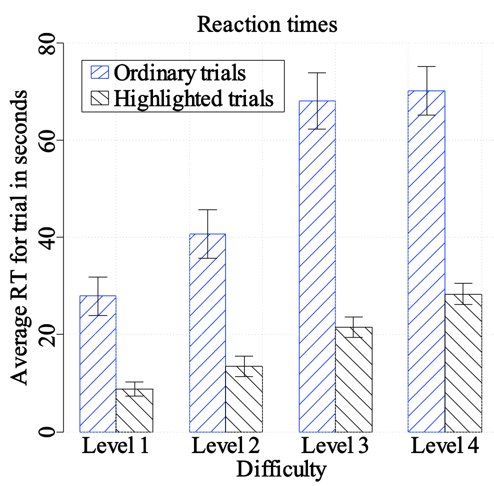

Figure 3. The mean reaction times in ordinary and highlighted trials clustered by the levels and averaged over all subjects.

Dimension Reduction

If subjects, indeed, used the dimension reduction strategy then the data should contain patterns of eye movements characterized by consecutive fixations on the cards sharing at least one common attribute value. To explore the existence of such pattern, the fixations from each trial were transformed into labeled fixation sequences. Each card in the round was assigned one area of interest with four different labels (see Figure 4).

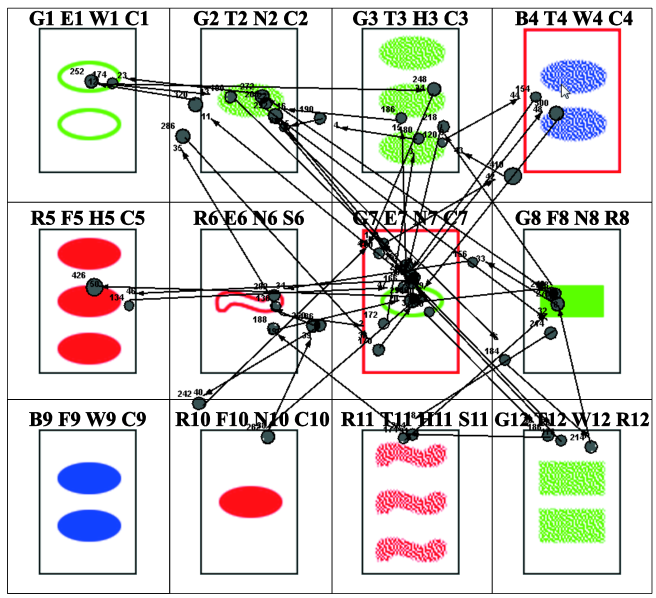

Figure 4. Array of 12 cards from problem “lvl3_15”. The highlighted card in this problem is card 7. Also shown are the fixations (circles) and saccades (arrows) produced by the subject. The outer thin black borders indicate 12 areas of interest. The four combinations of letters and numbers on top of each card represent four labels for each of areas of interest.

Each label described one of the attribute values in the card and the position of the card. Then each fixation was tagged with the labels of a corresponding area of interest within which it falls. The consecutive fixations on the same area of interest are considered as a single fixation, and the corresponding fixation durations are summed. Combining all labeled fixations with a common attribute type into fixation sequences produces four distinct sequences for each trial.

An analysis of the fixation sequences revealed the existence of a pattern of fixations related to the usage of the dimension reduction. We will demonstrate this using the example problem in Figure 4, which shows the saccade sequence of a particular subject finding a level 3 set (formed by the forth, fifth and seventh cards).

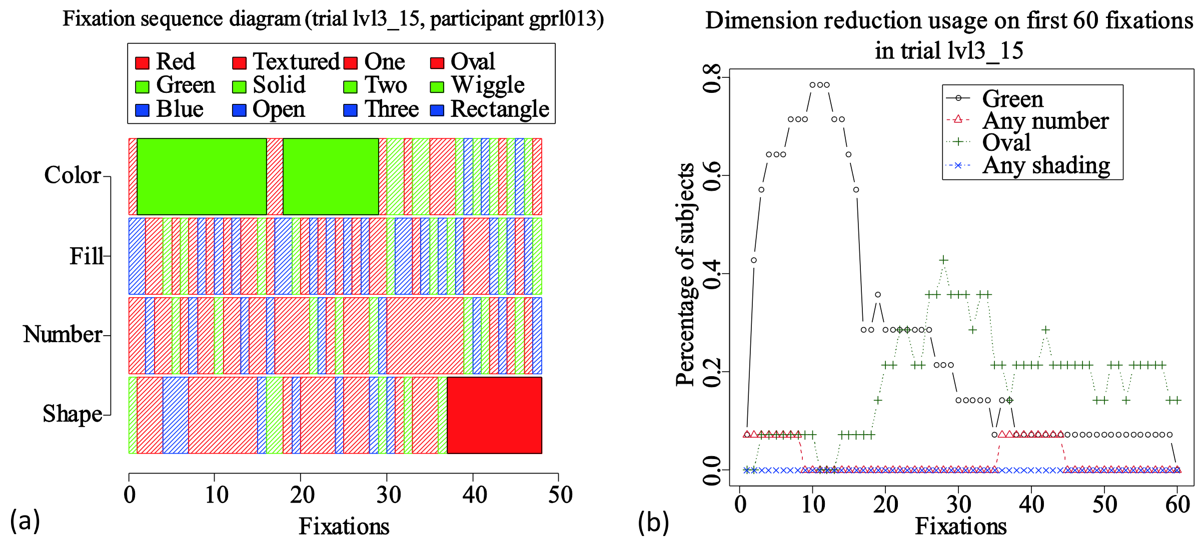

Figure 5. (a) Single subject's fixation sequence diagram for trial "lvl3_15"; (b) Changing proportion of subjects who used dimension reduction in trial "lvl3_15" as a function of fixation position in the sequence and attribute value.

Figure 5.a shows the fixation sequence diagram for that problem. Each lane in the diagram shows the subject's fixation sequence with respect to the particular attribute type. One unit on the x-axis represents a fixation on one particular card, while the corresponding bars on four lanes represent the attribute values of that card. In the diagram the labels are color coded according to the corresponding attribute value. The consecutive fixations on the cards with the same attribute value are shaded with a solid color if the probability of such a fixation

subsequence occurring by chance is equal to or below 0.05 (refer to Appendix for details of calculating the probability).

From the figure we can see that at the beginning of the trial, the subject looked at green cards and by the end at cards with an oval shape. This fixation pattern is deliberate since the probability of such pattern occurring by chance is below 0.05. It is possible that the subject tried to find a set with cards with one symbol in between, but this sequence was not long enough to be significant. As a result, it is possible to conclude that the subject used the dimension reduction strategy at least two times and each time with respect to the different attribute value: green and oval consecutively. The fixation pattern for this trial is not unique for this particular subject. Figure 5.b shows the proportion of subjects who used the dimension reduction with respect to the green, oval, any number or any shading values. The figure shows that at the start of the trial most participants preferred to search for a set among green cards and later switched to a group of cards with an oval shape.

Effects of an attribute type on dimension reduction. According to Jacob and Hochstein (2008), dimension reduction primarily occurs with the most abundant value. However, it can be observed from Figure 5.b that majority of the subjects prefer group of green cards to a group of cards with oval shape despite the fact that the latter is the most abundant value. As we have discovered, type of the attribute plays important role in deciding the value to be used for dimension reduction.

To find an effect of an attribute type, we have calculated an average proportion of fixation sequences where subjects used dimension reduction for all problems. The result indicates that blocks of fixations with the same attribute value occupy on average 46% and 35% of an overall fixation sequence in trials with and without highlighted card respectively. Note that these estimates are on the conservative side, because some sequences may have not been recognized because they cannot be distinguished from a random sequence. It is mainly because human subjects produce wandering fixations, fixations that have no particular purpose within the context of the task. Example of such fixation can be seen in Figure 5: subject was scanning green cards, but for some reason fixated on two non-green cards in the middle of scan. Such fixation can influence the calculation of the significance of the dimension reduction blocks.

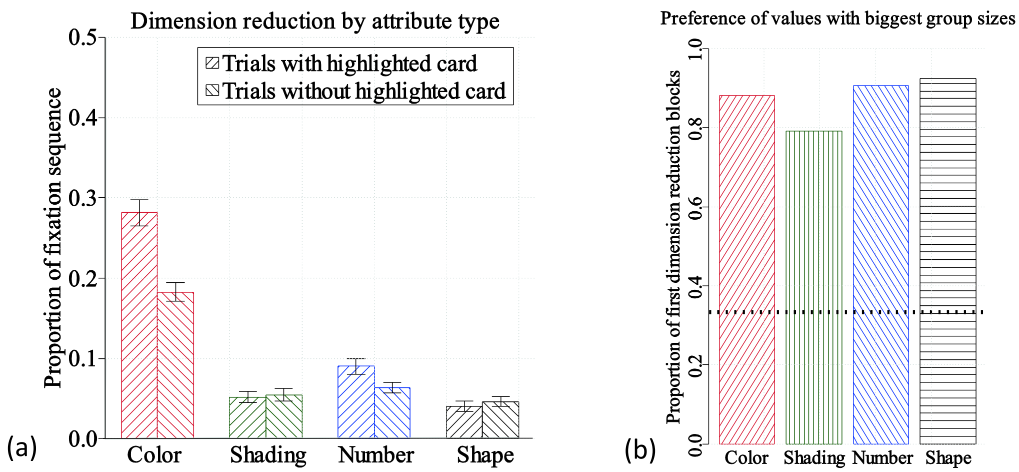

Figure 6. (a) Mean proportions of attribute types used in similarity-based scanning. Proportions are shown separately for trials with and without highlighted card. (b) Proportion of trials where subjects preferred to use for the dimension reduction a value with the biggest group size among other values of the same attribute type. The horizontal dashed black line indicates the expected proportion if the choice was made randomly.

Figure 6.a shows how the percentages distribute over the four attribute types and reveals an effect of subjects' attribute preference. Subjects are two times more likely to look at the group of cards with the same color than any other attribute. The results suggest that the four attribute

types might have different saliency properties with color being the most salient while shape and shading being the least salient attributes.

Effect of a group size on dimension reduction. To find an effect of a group size (Jacob and Hochstein, 2008) we considered each attribute type individually. If there is such effect then it should manifest itself among groups of cards with the same attribute type. For example, among groups of green, red and blue cards, the group with the biggest size should be chosen first for the dimension reduction.

For each attribute, we have calculated the proportion of trials where the value with the biggest group size (comparing to groups of the same attribute type) is chosen first for the dimension reduction. This proportion is then compared to the expected proportion if the choice was made randomly. To eliminate possible influence of the highlighted card, only trials without highlighted cards were considered.

As Figure 6.b indicates, in 85% of the trials subjects prefer the cards with a most abundant value within an attribute type. The trend is consistent among all four attribute.

Dissimilarity-based search

In the previous section we have seen that subjects use a dimension reduction strategy to reduce the complexity of finding a set. However, it is not yet clear how a similarity-based approach can eventually find sets with many different attribute values. For example, players cannot find level 4 set using dimension reduction strategy.

The fact that subjects were able to find level 4 sets given enough time suggests that the strategies they use is not limited to dimension reduction. In fact, Figure 6.a has already shown that subjects use dimension reduction strategy only 46% of the time, even though this number may be conservative given our analysis method. However, if we look at how the usage of the dimension reduction strategy changes during a trial, we see a clear decrease, as shown in Figure 7. Subjects prefer to use dimension reduction at the beginning of a trial and gradually stop using it as the trial progresses. It is our assumption that subjects gradually switch from a similarity-based strategy to a dissimilarity-based strategy. It should be possible to observe this switch from one strategy to another in fixation sequences produced from trials with highlighted cards.

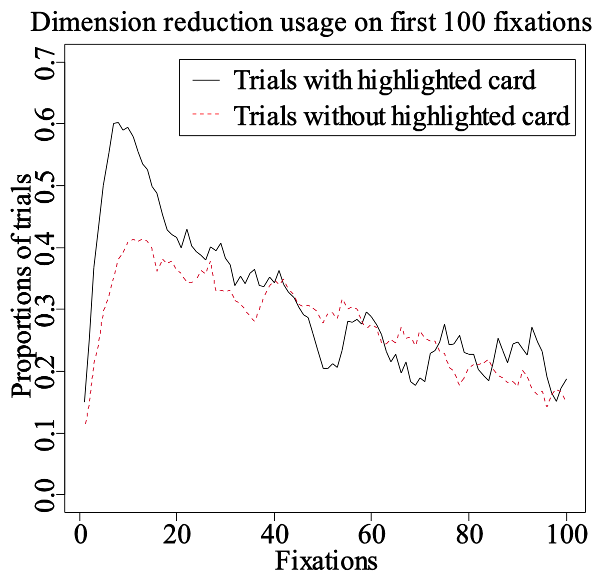

Figure 7. Changing proportion of trials in which dimension reduction was used. The proportions are calculated as a function of fixation position in the sequence and trial type.

Search subsequences. We analyzed the trials with one of the cards in the solution being highlighted. Preliminary inspection of the data revealed that subjects look back to that card approximately every five fixations, presumably to refresh their memory and to restart a new search subsequence. For example, the following fixation sequence for color attribute “R4

G7R11R10B3G7R2R11 R4 *B3R10R2B5B9B5B6* R4 *G7B5G8* R4”, with R4 being the fixation on the highlighted card, can be broken down into three subsequences.

Breaking down a trial into separate subsequences allows us to analyze the similarity between a highlighted card and the currently fixated card based on fixation’s position within a subsequence and position of the subsequence it belongs to. Figure 8.a shows how the average similarity of each individual card changes within a subsequence. Figure 8.b shows how the overall similarity of the cards in a subsequence changes from previous to next subsequences. The calculations were done separately for blocks of novice and expert players. There is a general tendency to look at a less similar card with each new fixation and each new subsequence. When players start the search they seem to prioritize the cards based on similarity to a highlighted card and look at more similar cards at the beginning of new subsequence (Figure 8.a). Furthermore, Figure 8.b suggests that with each new search subsequence the subjects lower the threshold for a similarity requirement and include in a visual search less similar cards that were not included in the previous subsequences. Finally, there may be a difference between expert and novice players in terms of bias to the similarity based search as Figure 8.b indicates. Expert players may be less biased toward similarity-based search than novice players.

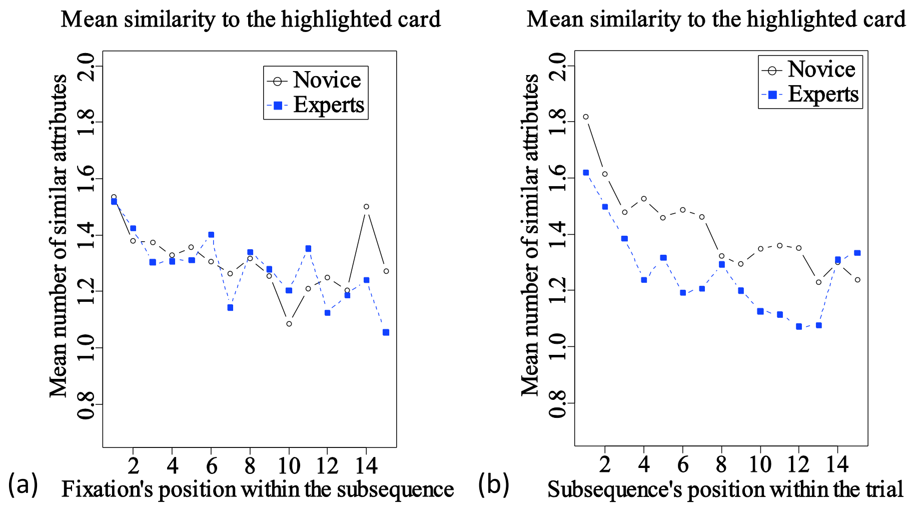

Figure 8. (a) The mean similarity of individual cards in a subsequence to a highlighted card. (b) The mean overall similarity of all cards in a particular subsequence to the highlighted card.

Using mixed-effect regression model (Baayen, Davidson & Bates, 2008), we have further investigated how the tendency to look at perceptually similar cards changes during the trial. The dependent variable in the regression model is the degree of a perceptual similarity (the number of same attribute values) of the next fixated card to the corresponding highlighted card (the values of Y axis in Figure 8). The following fixed effects were used in the model: *Subequence* is the log-transformed position of subsequence in the fixation sequence. *Fixation* is the log-transformed position of the fixation within the subsequence. The variable *Experience* represents the subject's level of expertise in playing SET (1 for novice group; 2 for intermediate group; 3 for expert group). In addition, two random effects on the intercept, *Subject* and *Trial*, were added to the model each representing subjects and trials respectively.

Table 1. The fixed effects' coefficients, t and p values.

Fixed Effects	Coefficients	Standard Errors	t values	p values
Intercept	2.028	0.079	25.592	0.000
Fixation	-0.177	0.016	-11.180	0.000
Subsequence	-0.162	0.011	-14.722	0.000
Experience	-0.065	0.022	-2.939	0.003
Fixation:Subsequence	0.034	0.007	4.684	0.000

Table 2. Variances and corresponding standard errors of the random effects.

Random Effects on Intercept	Variances	Standard Errors
Trial	0.104	0.322
Subject	0.004	0.062

The resulting coefficients for the fixed main and interaction effects are shown in Table 1. The table also presents the corresponding *t* and *p* values for the fixed effects. The variances and corresponding standard errors of the random effects are depicted in Table 2. In interpretation of the coefficients we are mainly interested in the signs. The positive coefficients increase the perceptual similarity to the highlighted card. Hence, the corresponding independent variables promote the similarity-based search. The negative coefficients decrease the perceptual similarity value. Therefore, the corresponding independent variables facilitate the transition from the similarity-based search to dissimilarity-based search.

Both *Fixation* and *Subsequence* have negative coefficients. This supports our assumption that over time cards that subjects look at decrease in similarity to the highlighted card. The fact that there are significant main and interaction effects of *Fixation* also indicates that transition occurs not only within fixation sequence as whole, but also within individual subsequences. The variable *Experience* represents the subject's level of expertise in playing SET. It has the significant negative coefficient. Both expert and novice players have initial bias to perceptual similarity-based search. However, expert players have smaller bias than novice players.

Effect of an experience on search pattern. As was discussed before there is a statistically significant difference between expert and novice players (Figure 8.b). However, so far the nature of searching behavior that results in this difference is unclear. To investigate this matter in more details we have analyzed the changing frequency with which subjects look at cards that have 3, 2, 1 or 0 attributes in common with the highlighted card. The results are shown in Figure 9.

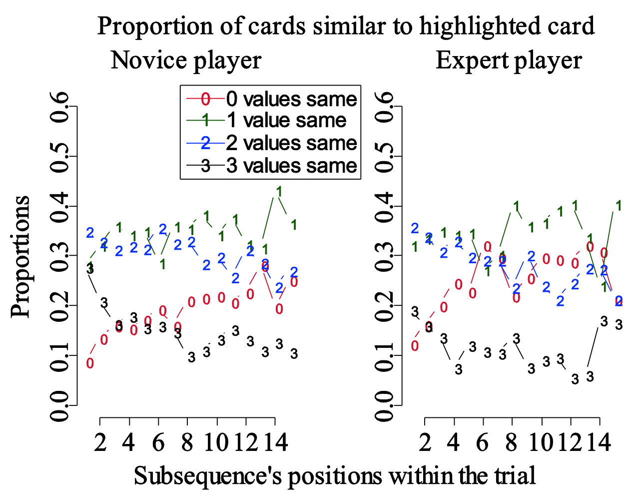

Figure 9. Graphs show how the perceptual similarity to the highlighted card changes over subsequences. The graphs depict the changing proportions of fixations on the cards that share 3, 2, 1 or 0 values with the highlighted card. The left and right graphs are for the blocks of the

novice and expert players respectively.

Next, four regression models were used to test the difference between each pair of lines of the same color. The models were used to test whether the slopes and the intercepts of two lines significantly differ. The first 14 subsequences from each trial of every subject were used in regression analysis. Each regression model has two main effects and the interaction effect: Subsequence (position of the subsequence within the trial), Experience (0 for novice players and 1 for expert players) and the interaction between the two. Note that the data from block of intermediate players was excluded from this analysis since we are interested in exploring the difference between two extremes of players.

Table 3. Coefficients for predictors in regression lines.

	Model 1 - Regression line for zero similarity			**Model 2 - Regression line for three similarity**		
	Coefficients	***t* values**	***p* values**	**Coefficients**	**t values**	***p* values**
Intercept	0.079	5.764	0.000	0.251	17.364	0.000
Subsequence	0.011	6.067	0.000	-0.010	-4.873	0.000
Experience	0.036	1.693	0.090	-0.045	-2.010	0.044
Subsequence *Experience	0.009	2.666	0.007	-0.001	-0.407	0.684
Degrees of freedom	1855					

We were mainly interested in significant main and interaction effects of Experience. Two models showed either significant main or interaction effects (Table 3). We can conclude that the major difference between expert and novice players is defined by their preference to search among the cards that are either the most (three attributes in common) and the least (none of the attributes in common) similar to the highlighted card.

Experiment Discussion

Experiment results' summary

Both the descriptive and mixed-effect regression analysis of the fixation sequences indicate that the subjects' basic strategy of playing SET is similarity based. Subjects prefer to look for a set among the cards that are similar to each other.

One specific instance of similarity-based strategy is the dimension reduction strategy (Jacob & Hochstein, 2008). The dimension reduction strategy can be used more than once (Figure 5 & Figure 6) within the same trial and each time with different attribute value. The player chooses one attribute value, to which we refer as a guiding value, and starts looking for a set among the cards that share that value. If the player fails to find a set with the current value, then another guiding value is chosen, and the new group of cards is defined as the next search space.

The overall strategy of dimension reduction is top-down, but the choice of a guiding value is heavily influenced by two bottom-up elements: (1) the size of the group of cards that share the value and (2) its attribute type. The importance of group size (Figure 6.b) was also found by Jacob and Hochstein (2008). However, contrary to their conclusion, we have found that the attribute type also plays an important role (Figure 6.a) in choosing a guiding value. Particularly, color is preferred to any other attribute type, while shape and shading are the least preferred attribute types. This result coincides with other studies concluding that people prefer to operate on colors rather than on shapes (Kieras, 2010; Kim and Cave, 1995; Pomplun et al., 2001). The number attribute also seems to be preferred to shape and shading, at least in trials with highlighted cards. The presence of a highlighted card can bias players to the values of that card. The bias can be so strong that it can override the effect of a group size or even attribute value during the choice of a guiding value.

Another interesting finding is the fact that within a trial subjects decrease the use of the dimension reduction strategy. This reduction (Figure 7) nicely coincides with gradual reduction in reliance on similarity (Figure 8, Figure 9 and Table 1). As the game progresses, players increasingly look at more dissimilar cards more suitable for finding higher level sets.

It seems that all players, independently of their experience, follow more or less these strategies. However, there are subtle differences between two types of players. We found that expert players are less dependent on similarity than novice players (Figure 8.b, Figure 9 and Table 1). Expert players are initially less likely to use dimension reduction and switch faster to the dissimilarity-based search than novice players.

A number of other results from the experiments indicate that expert players are less dependent on perceptual stimuli than novice players. For example, expert players do not need to refixate on the highlighted card as frequently as novice players do. In another example, more experienced players in required fewer fixations, hence, less number of perceptual stimuli, to compare three cards and decide whether they make a set (pattern of six or more repetitive fixations on same three cards followed by mouse clicks).

Additional assumptions

There are still open questions that were not answered by data analysis. For creating a plausible ACT-R model of SET player it is essential that we have complete picture of a player's behavior. In this section we address the essential but missing aspects of SET player's strategy by making our own assumptions or referring to relevant literature.

Choice of the guiding value. As was discussed earlier, the choice of a guiding value is not a completely conscious and heavily depends on bottom-up properties of the cards. We propose that the choice is defined by two components: task-independent factors that define saliency of an attribute value and task-dependent factors that define relevancy of a value to the current goal of a task.

The task-independent factors include attribute type and the group size. The four attributes have different inherent saliency properties. The color is the most salient attribute type, and the number is more salient than shape or shading (Kieras, 2010; Kim and Cave, 1995; Pomplun et al., 2001). On the other hand, six green cards are more salient than four red cards because of an effect of a group size on a saliency. These factors are not dependent on current goal and defined by the inherent properties of the visual object and a visual scene as whole.

The task-dependent factors include presence of a highlighted card and current progress of the trial. The task for the player is to find a set that includes the highlighted card (if it is present). This connection of the highlighted card to the current task increases the relevancy of the values in the highlighted card. On the other hand, if the highlighted card is green, and player is not able to find set among green cards then task-relevancy of the green value decreases.

It is our assumption that, in SET, frequently or recently used attribute values have a decreasing relevancy. For example, in the beginning of the game, most players tend to focus on the group of cards that share particular color or number values, since color and number are the most salient attribute types. However, its relevancies will decrease over the time, and eventually player will focus on other attribute types. Such mechanism gives a chance for the less salient attribute types, such as shape, to become a guiding value.

Dissimilarity-based search. A strategy of reducing the search space with a one value can also be used to find higher level sets. Let us assume that a player fails to find a set among cards that share the same color. In this case the player might choose, for example, one red card and look for the second and third cards among blue and green cards. Here the search space is again reduced since all but one card with a red value are ignored. Players may choose to use this only when dimension reduction strategy fails to find a set. The alternation between the dimension reduction and this strategy, with initial preference on former, can explain the gradual transition from similarity to dissimilarity.

Difference between novice and expert players. At this point one can be wondering if novice and expert players apply the same strategies then where the difference in reaction times comes from. It is our assumption that, although two types of players use same strategies, there are subtle differences between players in implementing these strategies. Particularly, expert players exercise more top-down control than novice players.

As it was discussed earlier, novice players need more frequent perceptual inputs than expert players. This is possibly why players prefer to use similarity-based perceptual processes to find sets than more top-down processes (Jacob & Hochstein, 2008). Naturally, it is much easier and faster to use bottom-up processes to process frequent stream of perceptual stimuli.

On the other hand, expert players need perceptual input less often which leads to two assumptions: (1) expert players somehow compensate for a lack of perceptual input (2) expert players are less dependent on perceptual processes and, therefore, can utilize more top-down processes. We suggest that expert players prefer to operate on more conceptual/abstract features, such as abstract representation of cards and rules. Naturally, such manipulation requires top-down control. For example, let us consider the situation where player has to find the third card given two cards with oval and rectangular shapes. The novice player does not exactly know what the third shape should be and may start looking for the cards with any shape that is different from oval and rectangle. On the other hand, the expert player may use conceptual rules to exactly predict that shape in the third card should be a wiggle. In such manner the expert player can predict all the values of the third card and significantly decrease time and effort required to compare the three cards.

The manipulation of the abstract features requires more skills and training. For example, in order to predict the right shape, experience with that is needed. Further, the presence of the abstract features decreases the dependency on similarity-driven perceptual processes. This can explain why expert players are less biased to bottom-up similarity-based processes and quicker in switching to dissimilarity-based search. In contrast, novice players are not likely to use rule-based cognitive processes, such as prediction, since they are costlier than similarity-based perceptual processes.

What interesting is that both conceptual and perceptual sides of the learning process were proposed separately from each other by Taatgen et al. (2003) and Jacob and Hochstein (2008). Taatgen et al. suggested an ACT-R based model of a SET player that relies on making predictions about the third card given two previous cards. The proposed approach is top-down and relies on generating the abstract representation of the third card. This model of the player fits the behavior of the expert player we have described in this paper. On the other hand, Jacob and Hochstein argued for the bottom-up strategy where similarity-driven perceptual processes play a major role. As we see it, the two studies describe two different aspects of the processes that lay on the opposite ends of the single spectrum defined by the player’s experience. In this study we unified the both approaches in a new cognitive model of a SET player.

How can bottom-up perceptual similarity-based reasoning transform into top-down rule-based reasoning? The transition from bottom-up to top-down processing in a SET player should be possible through training induced learning. The learning as we propose consists of two main aspects: firstly, the specific instances of a SET rule need to be learned; secondly, a player should encounter or apply the same rule several times to reduce its cost of use. Other studies (Goldstone & Barsalou, 1998; Shanks & Darby, 1998) suggest that it is possible to generated rule-based knowledge from lower level features and processes.

To test whether the assumptions made in this discussion are actually plausible we have implemented a computer model of the human player for SET. The model design decision includes both experimental finding and the additional assumptions we made about players' behavior.

An ACT-R Model of a SET Player

Model Design Decisions

We have implemented the model using the ACT-R cognitive architecture (Anderson, 2007). ACT-R has been successfully used previously to model players' behaviors for various games (Lebiere & West, 1999; West, Lebiere & Bothell, 2006; Taatgen et al., 2003). In each trial, the model is presented with 12 cards. One card is always highlighted indicating that it belongs to a set. The model has to find the other two cards forming a set. The trials from the experiment were used to test the model.

General strategy. The model largely follows the strategies that we have deduced from the data (Figure 10). At first, the model focuses on the highlighted card, and, based on its attribute values, chooses a guiding value. If the chosen value is present in a highlighted card then the model uses dimension reduction strategy otherwise dissimilarity-based strategy.

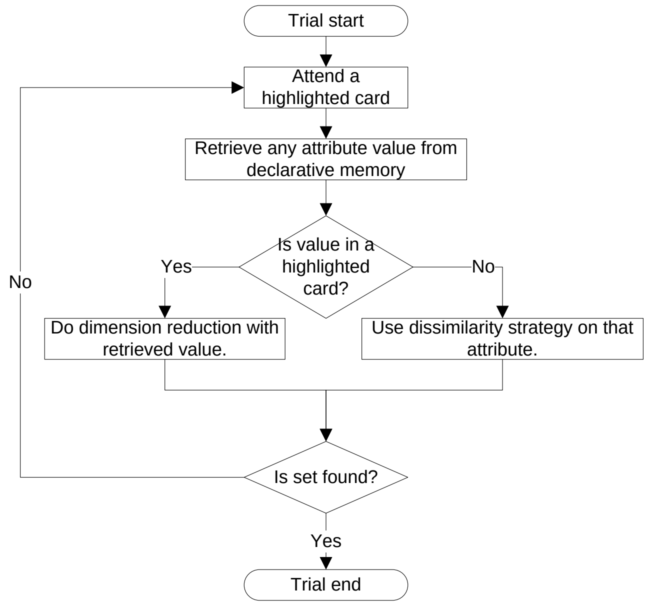

Figure 10. Model's strategy of playing set.

When using a dimension reduction strategy, the model looks at cards that contain the guiding value. When using a dissimilarity-based strategy, the model looks at cards for which the guiding value is different from that of the highlighted card. The implementation of the both strategies in the model is quite liberal in a sense that model behavior is not hardcoded. There is no explicit control over the guiding value choice. Neither there is an explicit top-down control over choice of which strategy to use and explicit timing of strategy change. The model decides all specific details of those steps by itself on the fly based on a visual scene and progress of the current trial.

Saliency and relevancy. The attribute value that is the most salient and relevant at the time is chosen as the guiding value. Saliency is a constant feature, however, relevancy is not and is calculated each time a new guiding value needs to be chosen. Within the model we have used ACT-R's activation mechanism to mimic both saliency and relevancy: the more activation the value has the more salient and relevant it is. Activation depends on several parameters such as values of a highlighted card, number of times the attribute value was used previously, the last time it was used, etc.

Two main parameters defining the saliency are attribute type and the size of the group of cards that have that value. The color is generally most salient attribute type followed by the number, while the shape and shading are the least salient types. The increasing logarithmic function was used to model the effect of a group size on a saliency.

The relevancy of a value depends on whether it appears on a highlighted card and whether it was used previously. The highlighted card spreads additional activation to each value it has. The relevancy of a value is temporarily inhibited after it has been used and no set was found. The time and duration of the inhibition are calculated according to Lebiere and Best's (2009) short-term inhibition equation. The complete description of the parameters used in calculating the activation is shown in Table 4.

Table 4. Parameters for calculating activation for an attribute value *i*.

Parameter	Influence	Implementation method
Attribute type	positive	ACT-R's equation for the base-level activation: $$B_i = \ln\left(\sum_{j=1}^{n} t_j^{-0.5}\right)$$ An initial number of references (n) is set for each attribute type as following (higher number result in higher activation): • Color chunks: 40 • Number chunks: 36 • Shape chunks: 32 • Shading chunks: 28 An exact calculation was used with the decay rate of base-level learning set to 0.5.
Group size	positive	Custom extension for ACT-R that spreads activation from a visual

		scene (visicon) to the declarative memory. The association weight parameter is set to 0.7. $G_i = \sum_k 0.7 * \ln(1 + fan_{ki})$
Highlighted card	positive	ACT-R's equation for a spreading activation from a visual buffer with *:visual-activation* parameter set to 0.8 and *:mas* to 4. $S_i = \sum_{\mathrm{j}} \frac{0.8}{6} (4 - \ln(fan_{\mathrm{ji}}))$
Frequency of use	negative (inhibitive effect)	ACT-R extension for a base-level inhibition (Lebiere & Best, 2009) is used with short-term decay rate and time scaling parameters set to 1 and 10 respectively. $I_i = \log(1 + (\frac{t_n}{10})^{-1})$
Latency of use	negative (inhibitive effect)	
Random noise	positive	ε_i – ACT-R's transient noise generated from logistic distribution with mean 0 and with *:ans* parameter set to 0.1.

Combining all parameters from Table 4 results in following equation for calculating activation for attribute value *i*: $A_i = B_i + S_i + G_i - I_i + \varepsilon_i$. This activation value is then used for calculating the probability of choosing value *i* relative to the other values via ACT-R's equation: $probability_i = \frac{e^{\frac{-A_i}{s}}}{\sum_j e^{\frac{-A_j}{s}}}$. At the end the value with the highest probability is chosen.

Top-down versus bottom-up. After deciding which strategy to use, the model proceeds with implementing that strategy. The search steps are the same for both strategies as it is shown in Figure 11. The model does the search in consecutive scanning cycles. In each cycle it tries to choose two other cards and check whether the three cards make a set. The model is free to choose its own scan path with the only restriction that it will not refixate on the cards it fixated before. The second and the third cards are usually chosen based on the order of the fixations: the earlier fixated card will be preferred. However, the choice of the third card can be explicitly controlled by a top-down process depending on the experience of the model.

Two different approaches are used in parallel to make the decision about the third card:

bottom-up and top-down. In the bottom-up approach the model searches for the third card solely based on the guiding value. The first card that is fixated by the visual thread and contains a guiding value is considered to be the third card. At the same time, the top-down approach tries to make a prediction about the third card based on the available rules. It generates the abstract representation of the third card and asks the visual thread to find the card matching that representation. The success and completeness of the prediction depends on the experience of the model. A more experienced model may be able to predict a whole card (all four attribute values), while less experienced model might even fail to make any predictions. More about experience and training is discussed in subsequent subsection.

Both approaches compete with each other. The approach that requires less time is favored to another. In other words, if the model is able to make a prediction before the visual thread fixates and encodes the third card then prediction is favored.

Given all three cards, the model verifies if the cards really make a set. If cards do not make a set then the model goes back to visual scanning. If a set is still not found then model interrupts the scanning and refixates on the highlighted card to choose another guiding value.

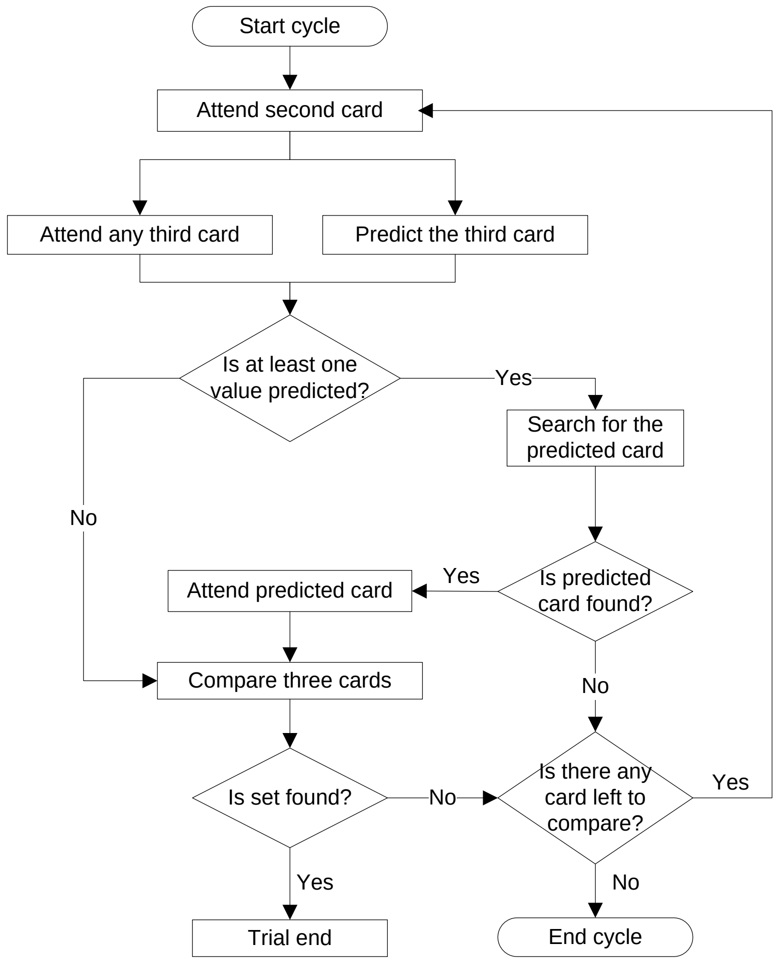

Figure 11. Model's strategy of playing set.

Experience and Learning. Both novice and expert models have abilities to predict the third card given two other cards. In other words, both models have necessary procedural knowledge to apply the rules. Rules themselves are declarative chunks and should not be confused with ACT-R's production rules.

The difference in experience between two models is defined by (1) a number of available rules and (2) previous repetition of those rules. The second factor defines the cost of using a rule. The more the rule is used the less its cost is. The cost of using a rule is defined by the time it takes to retrieve the rule from the long-term memory of the model. This time decreases as the rule is used more often. The decreasing time cost over period of training makes possible the competition between bottom-up and top-down processes described in the previous subsection.

A novice model has no rules at all. Therefore, it has to rely completely on bottom-up approach to find the third card. The expert model, on the other hand, was given enough training to generate all the rules and to minimize the cost of using them. Accordingly, it could use either bottom-up or top-down approaches depending on circumstances.

The model incorporates the training induces learning mechanism. After each trial, if a set is found, the model generates a random rule taken from the set it found. An example of such rule is Given(Textured, Solid) => Expected(Open). It should be noted that Given(Solid, Textured) => Expected(Open) and the previous rule are treated as different ones. In total, the model has to learn 36 rules: nine rules for each attribute. As the model plays, it gradually accumulates the rules. However, the rules need to be repeated frequently in order to be applied efficiently. Each rule is assigned an activation value indicating whether model can retrieve that rule from the memory, and how long the retrieval processes should take. Smaller values result in longer retrieval time or even in retrieval failure. Therefore, when the rule is generated for the first time, there is no guarantee that the model will be able to use it right away due low activation value. However, as the same rule is generated several times, its activation value increases, and the retrieval time decreases.

It happens that with enough rehearsal it takes little time to retrieve rules. Correspondingly, the top-down approach can make predictions before the bottom-up approach finds potential third card. As a consequence, with more rehearsal the model starts relying more on top-down approach than on bottom-up approach to find the third card. Such learning is gradual and can be generalized to SET problems on which the model was never trained before.

Threads. The model consists of two parallel processes (threads; see Salvucci & Taatgen, 2008) reflecting the top-downand bottom-up nature of the task. The bottom-up thread is responsible for visual processes such as deciding the visual scan path or shifting attention from one card to another. The top-down thread is responsible for the higher-level processes such as deciding the guiding value and comparing cards. Both threads can influence each other's processes indirectly. For example, the top-down thread chooses a guiding value based on what has already been tried earlier in the trial. However, bottom-up features such as what cards are visible or which card is being fixated also influence the choice.

Results

The resulting model was tested in novice and expert modes. In both modes the model had to play through 10 blocks where each block consisted of 30 trials with highlighted cards.

Reaction times. In Figure 12, the model's mean reaction times (dashed lines) are compared to the corresponding mean reaction times of human subject (solid lines). As it can be seen, the model closely reproduces the reaction times of both novice and expert human players. Similar to human players, the model also shows the tendency to have increasing reaction times with increasing difficulty of the set. Although model is very good at reproducing human reaction times, comparison at behavioral level is necessary for better evaluation of the model.

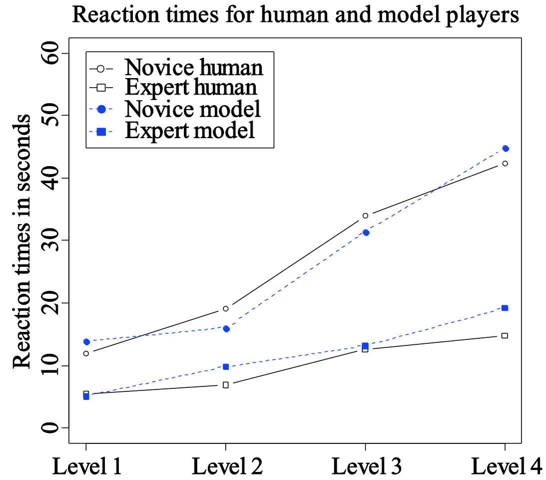

Figure 12. Reaction times of the novice and expert models comparing to the reaction times of the human players.

Dimension reduction. Both expert and novice models are quite good at replicating subjects' tendency to use dimension reduction and preference to certain attribute values. As an example, the expert model's fixation sequence in trial "lvl3_15" (Figure 13) is compared to that of human subjects (Figure 5). The model's fixation sequence closely resembles the sequence produced by the subject. At the beginning of the trial, the model also preferred to look at the green cards and later on switched attention to the group of cards with oval shape in a same manner as human subjects did. Analysis of the model's value preference over all blocks shows there is, in fact, intense competition between groups of green card and cards with oval shape (Figure 13.b). Over time, the model looses interest in green cards and mostly fixates on cards with oval shape. This decrease is consistent with behavior of the human subjects.

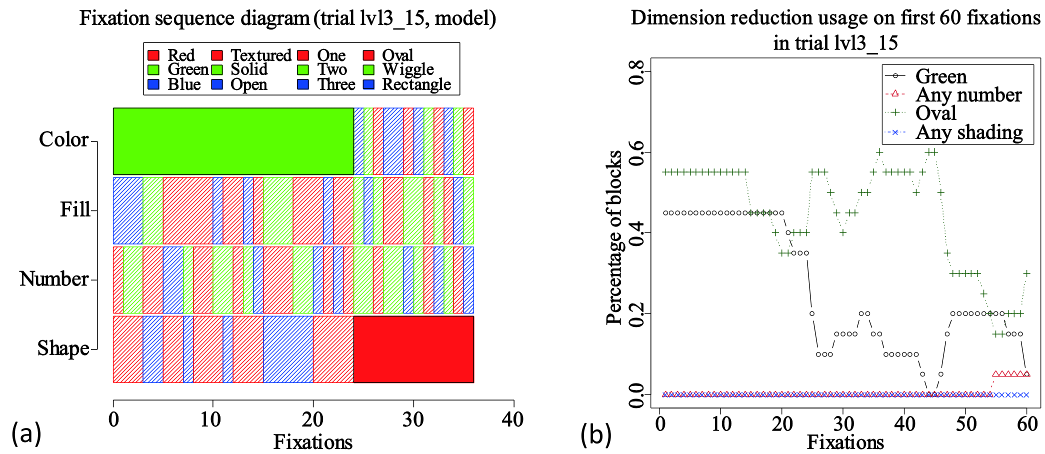

Figure 13. (a) Model’s fixation sequence diagram for trial “lvl3_15”. (b) Changing proportion of blocks where model used dimension reduction in trial “lvl3_15”. Proportions are shown as a function of fixation position in the sequence and attribute value.

Half of the times the model prefers to look at the green cards in the beginning of the trial although they form the second largest group after cards with an oval shape. Nevertheless, the fact that color is the most salient attribute type is enough to compensate for a smaller group size. Defining separate saliency values for attribute types works quite well for modeling players’ bias to an attribute types.

It can be observed from Figure 13 that the model favored shape in the later stage of the game, which is the least salient attribute type. This is due to the effect of a group size. Oval shape compensates its inherent lack of saliency with bigger number of occurrences. The fact oval value provides strong competition to green value even at the beginning of the trial suggests that the effect of a group size is stronger than it should be.

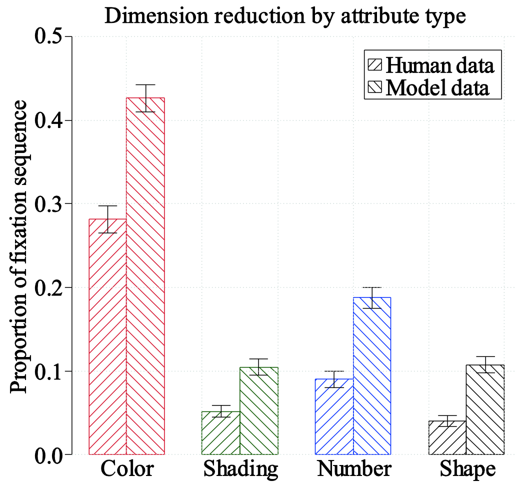

Figure 14. Mean proportions of attribute types used in similarity-based scanning. The overall values of all subjects' trials with the highlighted card are compared to the overall values of both models' trials.

Overall, the saliency and relevancy mechanisms work well in modeling subjects' strategy to use dimension reduction. Combined data from both models shows similar order of preference for the attribute types as the human subjects. Figure 14 shows that, in general, the models clearly prefer color and number while they make little difference between shape and shading. Both models gradually stop using dimension reduction if it fails to find a set (Figure 15). This behavior is again consistent with behavior of human subjects. However, models are more dependent on dimension reduction strategy than the human subjects. We attribute this difference to the difference in manner of scanning between model and human subjects. We discussed earlier that human subjects can get distracted and produce wandering fixations in the middle of the scan. On the other hand, model is precise and does not produce such fixations.

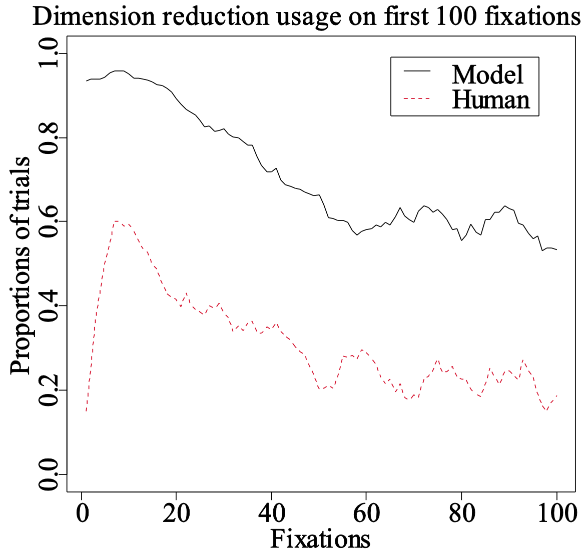

Figure 15. Changing proportion of trials in which dimension reduction was used. The proportions are calculated as a function of fixation position in the sequence. Model data include data from both expert and novice models.

Dissimilarity-based search. Our experiment revealed that the subjects gradually switch from dimension reduction strategy to a dissimilarity based search (Figures 8 and Figure 9). To test whether the model exhibits the same pattern of behavior as the human players, the same type of analysis was done on fixations sequences produce by the model. The results can be observed in figures 16 and 17.

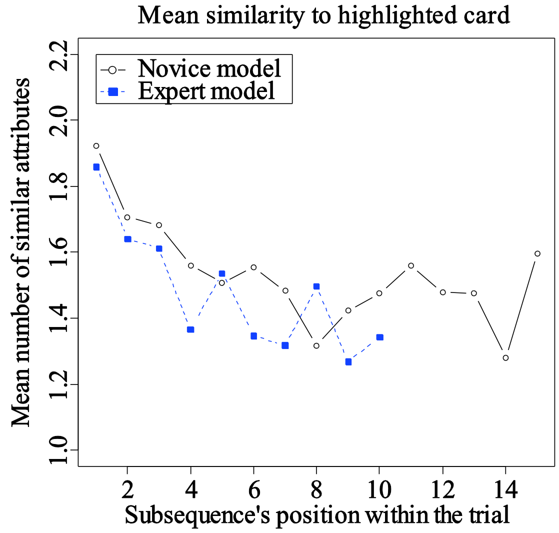

Figure 16. The mean perceptual similarity of cards in a particular subsequence to the highlighted card. Values are calculated from model's fixation sequences.

It is quite obvious that there are gradual transitions from the similarity-based search to the dissimilarity-based search for both expert and novice models. In general, the model's behavior corresponds well to the human player's behavior. However there are some differences. The model is more biased to perceptual similarity, otherwise dimension reduction, than the human players. As a result, the difference between novice and expert models with respect to initial bias to the perceptual similarity is smaller. The major difference between the expert and the novice models comes from the transition speed. It can be seen that graph for the expert model comes to an abrupt end at 10th subsequence. This is due to the fact that the expert model rarely required more than 10 subsequences to find the set. Nevertheless, the model demonstrates similar transition from similarity- to dissimilarity-based search as the human players.

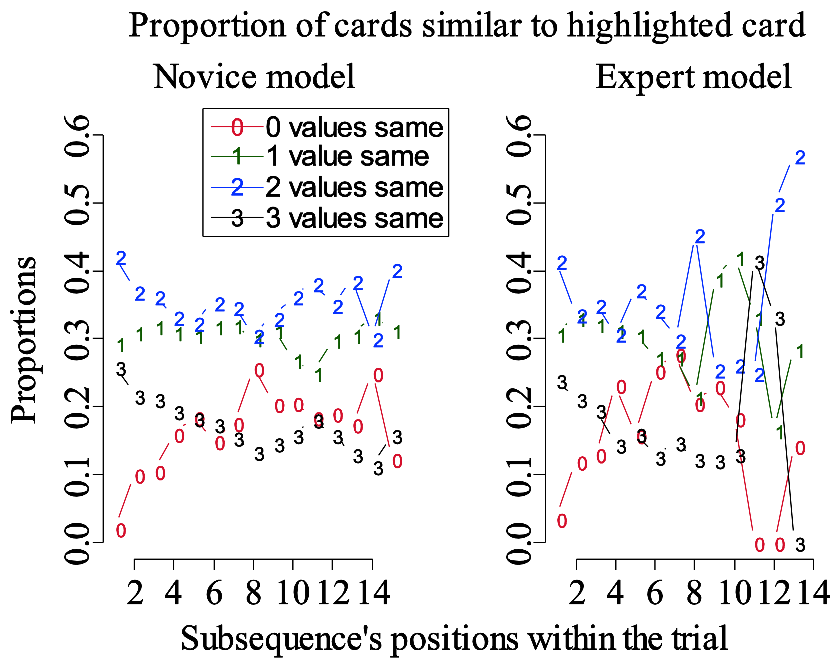

Figure 17. Graphs show how the perceptual similarity to the highlighted card changes over subsequences. The graphs depict the changing proportions of fixations on the cards that share 3, 2, 1 or 0 values with the highlighted card. The left and right graphs are for the blocks of the novice and expert players respectively.

General Discussion

Bottom-up and top-down processes

Improvement in playing SET comes from finding a proper balance between the two types of processes. Novice players initially tend to rely on bottom-up processes, because they lack the necessary skills and strategies. Improvement in the game is characterized by an increase in the involvement of top-down processes.

A similar development was found in studies of other games such as Scrabble (Halpern & Wai, 2007). In that study, novice and expert players also differed in the interplay between top-down and bottom-up processes. Novice players prefer to rotate and rearrange the letters physically to check whether they form a word. It makes players very much dependent on bottom-up motor processes and perceptual stimuli representing the letters. On the other hand, expert players prefer to rotate and rearrange the letters mentally. Hence, an expert player prefers to use top-down processes to manipulate the abstract representations of the letters.

Another example of a shift in balance between bottom-up and top-down processes is observed in Tetris. Initially it was believed that novice players prefer to rotate the tokens mentally to check whether that piece will fit at various parts of the screen while more experienced players prefer to rapidly rotate the tokens physically and rely on actual perceptual images of the tokens (Kirsh & Maglio, 1994). The reason, according to Kirsh and Maglio, is that physical rotation is more efficient than mental rotation. However, a later study showed that players with lots of experience prefer to rotate pieces mentally rather than physically (Destefano & Gray, 2011). This indicates that learning in Tetris happens also in form of transition from bottom-up processes to top-down.

In light of these findings, we conclude that such shift in balance between top-down and bottom-up processes may be a very common learning process.

Expertise

As it was discussed earlier on examples of SET, Scrabble and Tetris, the same task can be implemented with either top-down or bottom-up processes. However, the two types of processes will differ in amount of time and effort they require, and player can optimize the performance by choosing the one with the least cost.

Here one can argue that bottom-up processes are naturally faster and require less effort than top-down processes. However, (1) the cost of top-down processes can be reduced through a training induced learning, and (2) top-down processes give flexibility over a choice of the strategies that can be applied to accomplish the task. At some point two types of processes can become equally cost-effective, and with further training top-down processes might become more beneficial than bottom-up processes.

SET as an instance of real-world task

Jacob and Hochstein (2008) described SET as an instance of a categorization task. However, we would like to take even more general position. SET can be viewed as an instance of any real-world complex task in which either goal is not absolute, but changes with time or the environment is dynamic and unpredictable. In such tasks the absolute planning is impossible. Instead, one has to constantly change the plan (planning on the fly) to adapt to the new conditions and requirements. In this context the bottom-up reactive behavior to the stimulus from the external world is as important as the planning itself.

Let us consider SET as an instance of a task with changing environment and goals. In SET the array of 12 cards is not predetermined, and the player cannot make any detailed planning unless he or she has seen the cards. Even after looking at all cards it is impossible for the player to have exact plan since it is highly unlikely that the player will be able to memorize every value of every card. This is the unpredictable environment of the SET. Next, the ultimate goal for the player is to find a set, particularly any set. However, as it was shown, players prefer to search for the lower level sets at first, and, if not successful, they will start searching for higher level sets. This is an example of changing short term goal or subgoal. The feedback triggering the change of the subgoal is the time. If the player cannot find lower level sets for a long period of time then he or she will start looking for more difficult sets.

So how does SET relate to the complex real-world tasks? As an example we would like to refer to the previous study of modeling driving behavior (Salvucci, 2006). It happens that there common challenges in modeling a SET player and a driver.

Driving is very complex task that requires perceptual, motor and cognitive resources. What is even more important is that it is a safety-critical task where optimal visual attention allocation is crucial. Therefore, just like in SET, it is important to know which parts of the visual scene attract attention most. Needless to say that color, shape and shading play major role in defining regions of interest in a visual field.

Driving also includes both bottom-up and top-down processes. The amount of involvement of these processes depends on current goal and experience of the driver. For example, one of the tasks in driving includes keeping a car in a lane (check the visual world and steer the car if necessary). Drivers with little experience pay explicit top-down control over this task. However, for more experienced drivers the whole process becomes more automatic. They perform this task intuitively with direct motor reaction to the perceptual stimuli. Here is again transition from top-down to bottom-up processes.

The immediate goal can change often during a driving. One might switch from lane keeping to lane changing. The interesting point here is that driver switches from bottom-up processes that keep the car in the same lane to top-down processes. The driver has to explicitly make sure that there are no cars in the neighboring lanes before starting the maneuvers. This switch between the two types of processes is similar to behavior of SET player who uses bottom-up processes to find level 1 sets and top-down processes to find level 4 sets.

Conclusion

It is our hope to contribute to the understanding of visuo-spatial cognition where both internal conceptual knowledge and external perceptual stimuli converge in a goal-driven task.

As one step toward this goal we have studied the importance of perceptual and cognitive processes in complex tasks requiring both internal planning and reaction to perceptual stimulus from the environment. Our experiment and cognitive model show that both types of processes are involved in decision-making, and there is a complex interaction between them. In our model a major improvement in performance comes not from the optimization of one or another process, but from learning at the top-down level and finding an optimal balance between bottom-up and top-down processes.

We have briefly discussed how finding from this study can be generalized to other problem-solving tasks. Using examples of other games such as Tetris and Scrabble and more complex task such as driving we have described the similarities of behavior between those tasks.

So far we have investigated the conscious and unconscious perceptual influences on top-down planning. However, there is still an open question of how an application of bottom-up processes can benefit in reducing the cost of top-down processes. If a Scrabble player always manipulates letters physically then how can the cost of doing the same operations mentally are reduced? More studies are required to obtain complete picture of such learning processes.

Another important question is whether prior conceptual knowledge influences the perception of the external world. If a player likes green then will green cards be preferred over the red cards? Is such influence conscious or unconscious, and how can it be modeled? Those are the research questions that need to be investigated in the future.

References

Anderson, J. R. (2007). *How Can Human Mind Occur in the Physical Universe?* New York: Oxford University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-Effects Modeling With Crossed Random Effects for Subjects and Items. *Journal of Memory and Language , 59*, 390-412.

Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence , 47*, 139-159.

Clark, A. (1997). *Being There: Putting Brain, Body and World Together Again.* Cambridge, MA: MIT Press.

Destefano, M., Lindstedt, J., & Gray, W. (in press). Use of Complementary Actions Decreases with Expertise. *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp.). Austin, TX: Cognitive Science Society.

Goldstone, R. L., & Barsalou, L. W. (1998). Reuniting Perception and Conception. *Cognition , 65*, 231-262.

Green, C. S., & Bavelier, D. (2004). The Cognitive Neuroscience of Video Games. In P. Messaris, & L. Humphreys (Eds.), *Digital Media: Transformations in Human Communication.* New York: Peter Lang.

Halpern, D. F., & Wai, J. (2007). The World of Competitive Scrabble Novice and Expert Differences in Visuospatial and Verbal Abilities. *Journal of Experimental Psychology: Applied , 13* (2), 79-94.

Jacob, M., & Hochstein, S. (2008). Set Recognition as a Window to Perceptual and Cognitive Processes. *Perception & Psychophysics , 70* (7), 1165-1184.

Kieras, D. (2010). Modeling Visual Search of Displays of Many Objects: The Role of Differential Acuity and Fixation Memory. *Proceedings of the 10th International Conference on Cognitive Modeling*, (pp. 127-132).

Kim, M., & Cave, K. R. (1995). Spatial Attention in Visual Search for Features and Feature Conjunctions. *Psychological Science , 6* (6), 376-380.

Kirsh, D., & Maglio, P. (1994). On Distinguishing Epistemic from Pragmatic Action. *Cognitive Science , 18*, 513-549.

Kong, X., Schunn, C. D., & Wallstrom, G. L. (2010). High Regularities in Eye-Movement Patterns Reveal the Dynamics of the Visual Working Memory Allocation Mechanism. *Cognitive Science , 34* (2), 322-337.

Lebiere, C., & Best, B. J. (2009). Balancing Long-Term Reinforcement and Short-Term Inhibition. *Proceedings of the 31st Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Lebiere, C., & West, R. L. (1999). A dynamic ACT-R model of simple games. *Proceedings of the Twenty-first Conference of the Cognitive Science Society* (pp. 296-301). Mahwah, NJ: Erlbaum.

Pomplun, M., Sichelschmidt, L., Wagner, K., Clermont, T., Rickheit, G., & Ritter, H. (2001). Comparative Visual Search: A Difference That Makes a Difference. *Cognitive Science , 25* (1), 3-36.

Rayner, K. (1995). Eye Movements and Cognitive Processes in Reading, Visual Search, and Scene Perception. In J. M. Findlay, R. Walker, & R. W. Kentridge (Eds.), *Eye Movement Research: Mechanisms, processes and applications* (pp. 3-22). Amsterdam: North Holland.

Salvucci, D. D. (1999). *Mapping Eye Movements to Cognitive Processes*. Doctoral Dissertation, Carnegie Mellon University, Department of Computer Science.

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

Salvucci, D. D., & Taatgen, N. A. (2011). *The Multitasking Mind.* New-York: Oxford University Press, Inc.

Salvucci, D. D., & Taatgen, N. A. (2008). Threaded Cognition: An Integrated Theory of Concurrent Multitasking. *Psychological Review , 115* (1), 101-130.

Shanks, D. R., & Darby, R. J. (1998). Feature- and rule-based generalization in human associative learning. *Journal of Experimental Psychology: Animal Behavior Processes* , *24*, 405-415.

Taatgen, N. A., & Lee, F. J. (2003). Production Compilation: A Simple Mechanism to Model Complex Skill Acquisition. *Human Factors , 45* (1), 61-76.

Taatgen, N. A., Oploo, M., Braaksma, J., & Niemantsverdriet, J. (2003). How to Construct a Believable Opponent Using Cognitive Modeling in the Game of Set. *Proceedings of the Fifth International Conference on Cognitive Modeling*, (pp. 201-206).

West, R. L., Lebiere, C., & Bothell, D. J. (2006). Cognitive architectures, game playing, and human evolution. In R. Sun (Ed.), *Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation* (pp. 103-123). New-York: Cambridge University Press.

Footnotes

[1] SET is a game by Set Enterprises (www.setgame.com)

Appendix

The probability of *k* subsequent fixations falling on cards that share at least one value in common if the fixations are assumed to be random is calculated with a following equation:

$$P = \sum_{i=1}^{4} \sum_{j=1}^{3} \frac{n_{ij}}{11} * \left(\frac{n_{ij} - 1}{11}\right)^{k-1}$$

k – the number of fixation in fixation subsequence

n_{ij} – a number of cards in array of 12 cards that have value *j* for an attribute *i*

If the calculated probability of the block of k fixations is below 0.05 then it is assumed to be not produced by chance. The blocks are calculated for an each attribute type. If two blocks of fixations from different attributes overlap then the block with the least chance probability is preferred. The other block is cut at the point of an overlap, and its probability is calculated again based on the block's new length. If the two blocks overlap and have an equal chance probability then the longest block is preferred. If the lengths are also equal then one of the blocks is randomly chosen and removed. Finally, Holm-Bonferroni correction was used on initial significance value of 0.05. The correction compensated for the inflation of the chance probability when multiple solid blocks are present in same trial.