Online Appendix for "Predictable Effects of Visual Salience in Experimental Decisions and Games"

Xiaomin Li

Colin F. Camerer

A. Image Appendix

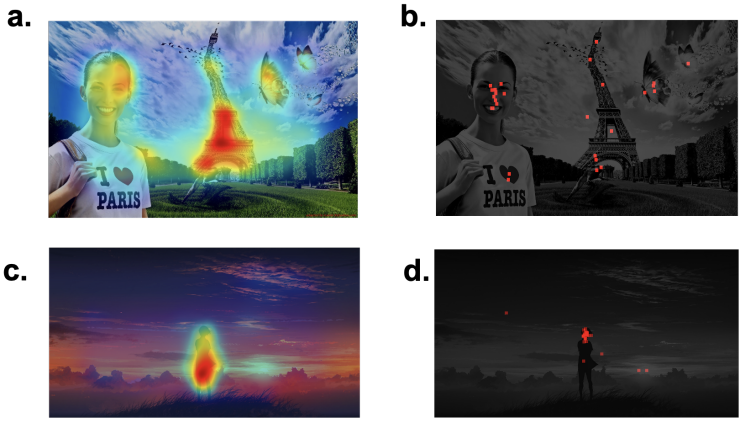


Figure A1. Examples of different salience centers

Note: (a) is an image with seven salience centers; (c) is an image with one salience center. (b,d) are corresponding maps (red dots) of actual choice data in each matching game. The choice map in (b) is more dispersed because the salience centers in (a) are more numerous.

B. History and details of saliency algorithms

The SAM algorithm takes one image as an input and outputs its predicted saliency map. The saliency map is a saliency value from zero to one (least salient to most salient) assigned to each pixel on an image. Figure I in the text is a specific example of the SAM saliency map from one of the pictures we used in our experiments. A little history of saliency mapping may be useful here to convey how well-founded these algorithms are.

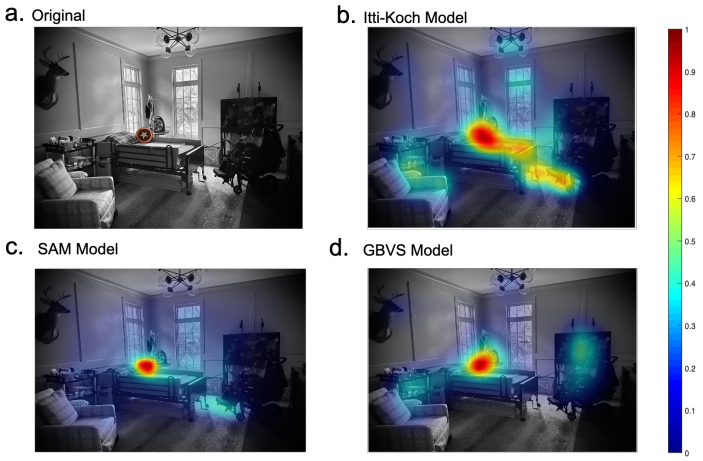


FIGURE B1. COMPARISONS BETWEEN THREE SALIENCY MODELS

Note: We show the result or different saliency model on the same image (a). Both the Itti-Koch model and GBVS are fully interpretable (shown in b and d). Also in this example GBVS (Harel et al., 2007) and SAM have very similar results. All three outputs are color-plot in the same standard colormap function in matlab (type: "jet").

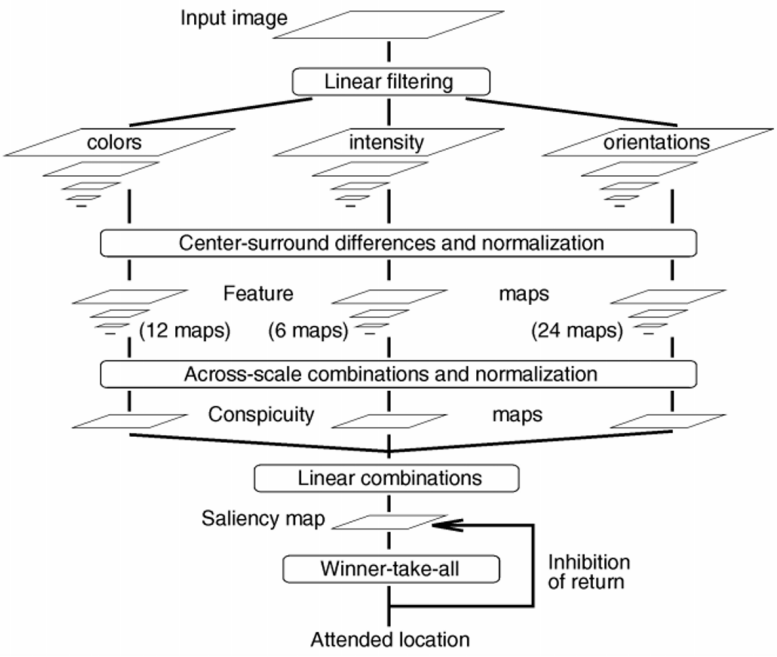
Inspired by a deep understanding of how the human visual system prioritizes attention, a series of progressively improving algorithms were developed to use visual images as inputs, and output predictions about where people will look in the first 1-2 sec of processing (Itti et al., 1998; Harel et al., 2007; Judd et al., 2009). Figure A1 compares two early algorithms and SAM in one example image.

Figure B2 shows an early algorithm from Itti et al. (1998). These early algorithms used a combination of handcrafted features to extract information about contrast, color, and orientation. Dark-light contrast is special because it marks boundaries between objects. Color and orientation are also thought to have adaptive value in parsing images in ways that are ecologically useful.

Consider the stick figure "I". The bottom-up perception is a black vertical line of a certain length, with slightly extended top and bottom horizontal lines on top of the vertical line, surrounded by contrast with a white background.

A more abstract high-level concept of bottom-up is that (depending on the

FIGURE B2. ITTI-KOCH MODEL



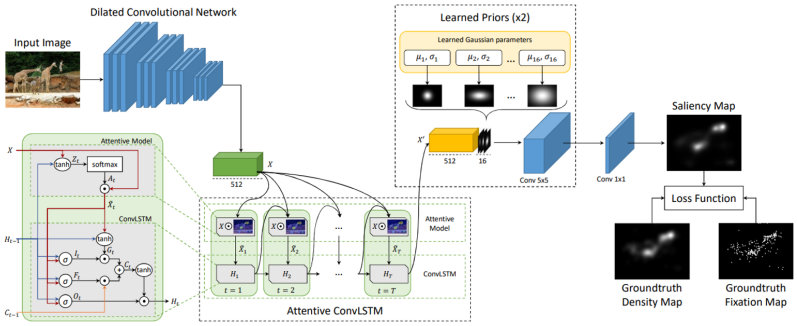
Note: presents the model of Itti-Koch (1998).

subject population) the I may be more familiar, valuable, or behaviorally useful. An English speaker will perceive "I" as a marker of first-person communication; a student just learning Roman numerals will perceive "I" as the number one; and an architecture aficionado may perceive it as an Iconic column, a part of a building. All the latter forms of salience use semantic knowledge– which is local and acculturated– about the world to inform the perception of what "I" means and what to do with that information. Which features or objects are personally relevant, valued, familiar, and novel will also be trained into a bottom-up algorithm, but will depend on the image set and characteristics of the subject population.[64]

The SAM algorithm (see Figure B3) we use was tuned using human free gaze

[64]See studies on the effects on perception of recent choice history (Awh et al., 2012), familiarity and novelty (Itti and Baldi, 2009), value for consumer goods (Towal et al., 2013), and self-reported "meaning" (Henderson and Hayes, 2017).

FIGURE B3. SAM MODEL



Note: presents the deep neural network model structure of the state-of-art SAM model framework.

TABLE B1—SUBJECTS AND SOURCE IMAGE INFORMATION OF SAM TRAINING

Dataset	**Subjects restriction**	**Image Categories**
Salicon	Age: 19-28	91 categories including persons, vehicle, outdoor, animal, and etc.
MIT1003	Age: 18-35, We guess from Boston/MIT area.[65]	natural indoor and outdoor scenes
MIT300	Age: 18-50 We guess from Boston/MIT area	natural indoor and outdoor scenes
CAT2000	age 18-27, observers were undergraduates at USC from different majors and from mixed ethnicities.	Action, Affective, Art, Black & White, Cartoon, Fractal, Indoor, Inverted, Jumbled, Line Drawing, Low Resolution, Noisy, Object, Outdoor Man-made, Outdoor Natural, Pattern, Random, Satellite, Sketch, and Social.

data on a large number of images, without any special goals or incentives. The subject are just told to look. These algorithms were **not** designed to predict active choices in games with specific goals, such as matching and hider-seeker games. The matching goal, for instance, is to choose a location another person is also likely to choose. This is a goal-directed influence on perception which is likely to produce visual fixations that are different from free viewing. Thus, the extent to which SAM can predict the influence of predicted salience is probably a lower bound on how well better models, incorporating top-down goals, will do.

To help clarify, the origins of SAM and previous algorithms, Table B1 describes

the sets of images used and some characteristics of the subjects whose free gaze data were used to train SAM. The original papers are not crystal clear on who the subjects were, which is an indication that the authors think of the perceptual processes they are studying as rather homogeneous across people.

C. Focality in previous game experiments

There is a substantial, interesting series of experimental studies about focality in matching games. These studies are quite different from our approach but are described here for completeness.

There was a long lag between Schelling's early 1960 discussion and later bursts of careful experimentation on focality.

Mehta et al. (1994b) proposed an important contrast between "secondary salience" and "Schelling salience". Following Lewis (1969, pp. 24-36), they suggested that when players are not sure what to choose, they choose according to "primary salience", which is "some (possibly stochastic) process that brings one of the labels to the player's mind" (p. 660). *Secondary* salience is the belief about what creates primary salience for *others*. This process can obviously be iterated further.

Their experiments supported this distinction. In "picking" conditions people just picked an object from a choice set (e.g., a set of flowers). In "matching" conditions their choices were matched with randomly chosen others and rewarded if they matched. The hypothesis is that picking measures primary salience and matching measures secondary salience. Indeed, the most common modal choice in the picking condition was usually chosen much more often when matching.

Note that this primary-secondary distinction is instantiated naturally in the SCH model (although that model was developed to explain behavior in a much wider range of games). In SCH, the process that brings one of the labels to the player's mind– its primary salience– is predicted *ex ante* from the bottom-up SAM model. In the Mehta et al. (1994b) paradigms primary salience has to be *measured* by having people choose objects in the picking condition. Using

SAM a primary salience prediction is delivered for all images; no new data or free parameters are needed.

In contrast to primary and secondary salience, an object has "Schelling salience" if it is unique or is chosen by a rule that leads to unambiguous results, which improves matching. Schelling salience need not arise from primary or secondary salience. For example, in a list of historical figures including Adolf Hitler, Hitler could be Schelling-salient even though few people would pick Hitler (primary salience) or think others would pick Hitler (secondary salience). Indeed, Mehta et al. (1994b,a) find evidence for both secondary and Schelling salience in their data.

More ambitiously, Bacharach (1993); Bacharach and Bernasconi (1997) proposed general principles underlying focality in matching choices from sets of objects, essentially trying to unpack Schelling salience into component parts. Their idea was that if people know their goal is coordination, they will try to naturally categorize objects into subsets and chose from more distinctive– e.g., smaller– subsets. However, subjects' actual choices were not always consistent with the most non-obvious of their principles. There experiments are elegant and careful. They were held back by the fact that a key element of the theory– "noticing" set-theoretic features– is measured only crudely (by self-report), whereas we now have eyetracking to measure noticing directly.

Focality is also likely to work differently in hider-seeker games (HS). Studies by Mehta et al. (1994b) Bacharach (1993); Bacharach and Bernasconi (1997) were focused on coordination; at that time in the research history, there was no ambition to create theories of focality that would span games of different competitive structures. Understanding matching was difficult enough.

In a separate strand of cumulated regularity, an early study by Rubinstein et al. (1997) (RTH) used a four-choice hider-seeker game. Their canonical example is a choice between four letters ordered from left-to-right, where one letter is a singleton subset, like so:

A B A A

RTH hypothesized that the left and right A letters are avoided (because of "extremity-aversion"; cf. Bar-Hillel 2015). They hypothesize that the single B is clearly focal because it is both visually and semantically unique, and it will therefore be avoided by hiders. That leaves the second "interior" A from the right, which is least focal when compared to other choices (and therefore uniquely non-focal, giving it an ironic strategic focality due to uniqueness).

In these early studies, extremity-aversion and B-focality are simply hypothesized intuitions; they were not guided by data or visual perception principles. On this basis, RTH predicted that the third A would be chosen most often. Indeed, in their experiments, the third A is chosen most frequently both by hiders (40%) and seekers (45%). As a result, there is a "seeker advantage" because the seekers win more often than Nash equilibrium prediction of 25%. However, our replications in Caltech and UCLA subjects found much lower rates of the choice of the third "inner" A, around 29%, closer to the Nash 25% prediction (unpublished data).

Falk et al. (2009) used visual hider-seeker games similar to the four-letter choice. One game required choosing 3 cells out of the 25 locations in a 5x5 matrix. They observe both an edge aversion and a seeker advantage.[66] There is a lot of other interesting data and psychology in their paper. In a recent study Brocas and Carrillo (2021) targeting at development and social choices, seeker advantages in hider-seeker games were discovered early in the life stages (second grade kids), and were still present in the adolescents control group (with effect size increasing with age).

In the main text we noted that our modeling builds upon Crawford and Iriberri (2007a) (hereafter CI), they advanced a novel analysis of games like ABAA, based on level-k modeling. They hypothesized that behavior could be consistent with

[66]Based on data reported in their paper, the seeking win rate in this experiment is 10.37% while the chance level is only 6.25%, implying a seeker advantage of +4.12%. These numbers are rather close to our own, which are about 7% and 9%, although the paradigms differ a lot.

a level-k approach [67], in which level-0 behavior is influenced by salience. Rather than using an algorithm to predict salience, salience is parameterized by the frequencies of the outer A's and the central A. CI also assumed that level-k types only best respond to level k-1 types and that the population didn't contain any actual level zero types. Under this framework, they estimated both level zero players' preferences towards different options (saliency biases) and population frequencies of level types. The general approach fits behavior well. Our paper expands on this approach by predicting saliency independently of choice, using no new data, in location games.

Hargreaves Heap et al. (2014) questioned the strength of the CI conclusions on the grounds that the salience of the extreme A's and the central A were estimated parametrically and not constrained across game structures. They created choice sets with a single "oddity" that is visually or semantically unique (e.g. a list of words which are all diseases plus the word "fitness".) They test whether the oddity is equally salient for level 0 players in three types of games– coordination (matching), discoordination (players win if they both choose something different), and hider-seeker. They reject the hypothesis that level 0 salience is the same across games. Crawford (2014) commented on their paper.

We find better "portability" of salience across matching and hider-seeker games. Specifically, we are able to predict the saliency-choice correspondence in matching games from SCH hider-seeker estimation.

D. Results from no-feedback trials

The realized matching rates when there is no feedback are as follows. Note that the hider-seeker matching rate (9%) is the same as in trials with feedback:

Figure D1 below is a quantile to quantile plot, plotting the percentage rank of saliency for each location against the percentage rank of choice frequencies for those locations in matching games. To get the Q-Q plot, we first mapped

[67] See Stahl II and Wilson 1994; Nagel 1995 and see Crawford et al. (2013) for a thorough review.

Table D1—realized matching rate

	No feedback	Number of observations
Nash mixed prediction	0.071	
Matching game	0.35(0.004)	550
Hider-seeker game	0.09(0.002)	523(s)+527(h)

Note: Statistical tests are against the null hypothesis that the seeker win rate is the baseline level and choices are independently and identically distributed across subjects (which is the Nash benchmark prediction).

all users' choice data (not only click points, but all points which fell into the circle) onto a one-dimensional saliency value, normalized from zero to one. (The highest saliency point in each entire image is one, and the lowest is zero). Then we ranked all these realized saliency values for all choices in the targeted sub-block. We also transformed the rank of the choice frequencies across all subjects into rank percentages. We plotted the normalized saliency value, which was also the percentage of saliency, against the percentage of points chosen with the same saliency ranking. The Q-Q plot below shows that all quantiles of choice data are above the same quantiles of saliency level, and hence above the diagonal dashed line that would result if people were choosing independently of saliency.

Figure D2 presents both Q-Q plots and density maps in the hider-seeker game. Figures D2 a-b indicate that seekers' choices are more biased towards salient locations than hiders' choices are, and both are much less saliency-biased than in the matching games (recall Figure D1). Keep in mind, however, that the hiders should be choosing locations as low in salience as they can perceive (i.e., a best-response Q-Q curve would be underneath the 45-degree identity line).

The density maps in Figures D2 c-d take every location in every game, and assign each one a saliency level (0-1 normalized within each image), and computes the frequency with which "strategies" (=locations) were chosen across all games and subjects. For hider-seeker games, these should be flat horizontal lines in equilibrium(except for sampling error). However, there are a disproportionate number of choices of high-saliency locations (that is, the densities turn up sharply at the right end of the scale). Seekers choose the highest-saliency locations about

three times as often, and hiders choose them about two times as often. There is a slightly disproportionate tendency to choose the lowest saliency locations (near zero at the left end of the scale), especially for hiders.

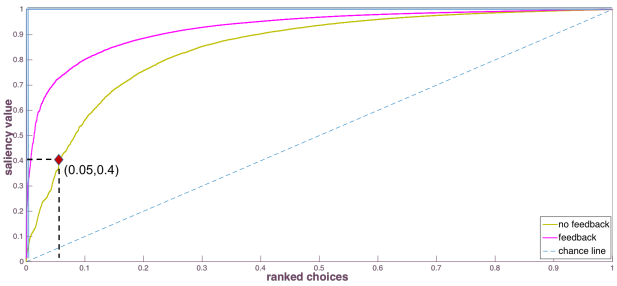


FIGURE D1. MATCHING GAME Q-Q PLOT OF CHOICE FREQUENCY(X-AXIS) AND SALIENCY RANKS (Y-AXIS)

Note: The red-diamond point (0.05,0.4) indicates that only 5 percent of choice points were made at the locations at or below 40% salience. Equivalently, 95% of the points fall within the top 60% most salient points. Choices generated by chance thus correspond to a diagonal line of this plot from $(0, 0)$ to $(1, 1)$. The maximal accuracy is the blue line: $y = 1$ for all $x > 0$, which would occur only if all choices fall on exactly the most salient point.

E. SCH MODEL COMPARISON WITH DIFFERENT SPECIFICATIONS

In this subsection, we are going to compare four different sub-models. We choose the Bayesian information criterion (BIC) to be the criterion for model selecting, since it balances the goodness of fit and the possibility of overfitting.

In all cases we restricted the softmax sensitivity parameter λ from 0 to 100. Larger values carry little extra information since $\lambda = 100$ is close to best response. Constraining λ also makes it easier to create a bootstrapped confidence interval, which is useful due to the non-smoothness of the target function (likelihood function).

Here are descriptions of models we are going to test (and see Table E1):

- Model 1: There are only two types of players: 1) naive players who play as level zero players described in the main text. 2) equilibrium players who do

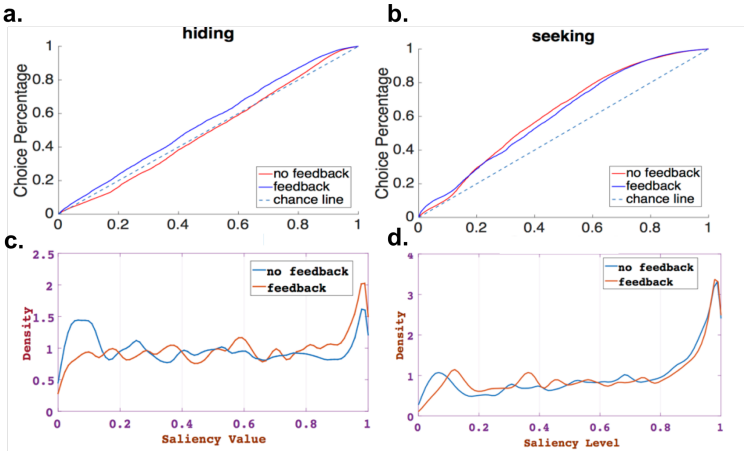


FIGURE D2. HIDER-SEEKER GAME Q-Q PLOT OF CHOICE FREQUENCY(X-AXIS) AND SALIENCY RANKS (Y-AXIS)

Note: a, b: Q-Q plots for hiding role (a) and seeking role (b). c, d: Kernel pdf density map of the choice frequency as a function of location saliency ranks. The x- axis is the rank of the saliency values and the y-axis is the probability density. Note: The kernel is Gaussian. The bandwidth is calculated using the formula: $\sigma \times \frac{4}{3N}^{0.2}$, in which σ is the standard deviation of the samples and N is the number of observations.

pure randomization. Both the proportion of naive players, p_s and p_h serve as free parameters.

- Model 2: There is no level zero player in the real population, but higher level types believe there is. The hiders and the seekers have different τs but the same μ and λ.

- Model 3: Same as model 2, except that level zero players exist both in the belief structure and in the population.

- Model 4: This model fits hiding data and seeking data separately using two sets of parameters. Each game has three parameters: μ, λ, and τ. The best fit model of it dominates the best fit of model 3 since model 3 is a special case of model 4. However, model 4 allows more free parameters, which the BIC value will penalize.

- Model 5: This model fits hiding data and seeking data using a common μ,

λ, but uses the level-k belief framework [68] rather than CH, assuming the population consists of players whose level ranging from one to four (no level 0's).

- Model 6: This is the same as model 5, except it allows level 0 types.

Table E1—Model comparisons for hider-seeker game

Model	Description	Free parameters (Estimated)	AIC	BIC
1	Level 0+equilibrium	p_s,p_h [1, .3]	12716	12728
2	Role-specific τ_x, f(0)=0	μ,λ,τ_s,τ_h [.004, 99, .46, .002]	12780	12803
3	Role-specific τ_x, f(0)$\neq$0	**μ,λ,τ_s,τ_h [.06, 100, .40, .10]**	**12650**	**12673**
4	Role-specific τ_x, μ_x, λ_x	μ_s,λ_s,τ_s,μ_h,λ_h,τ_h [.01, 90, .40, .07, 90, .50]	12646	12680
5	Level-k role-specific f(k), f(0)=0	μ,λ,$f_s(1)$,$f_s(2)$,$f_s(3)$,$f_h(1)$, $f_h(2)$,$f_h(3)$ [1, 99, .22, 0, .78, .83, .05, .12]	12681	12738
6	Level-k role-specific f(k), f(0)$\neq$0	**μ,λ,$f_s(0)$,$f_s(1)$,$f_s(2)$,$f_s(3)$, $f_h(0)$,$f_h(1)$, $f_h(2)$,$f_h(3)$ [.18, 99, .29 .05, 0, .66, .17, .22, .61, 0]**	**12652**	**12709**

Note: Each model in the table is specified in the text list. BIC is defined as -2*logL + numParam*log(numObs) and AIC is -2*logL + 2*numParam

Table E1 lists the best fit results of each model. Both BIC and AIC indicate that model 3 is the best performing model. Model 2 performs worst for the reason that without level zero types, the model structure will over predict the frequency of pure anti-salient hiders, which is not seen in the data..

The Level-k Model 6 is almost as accurate by AIC and BIC, and we commented on what can be learned from it in the text. Figure E1 plots predictions of that model and the data, for comparison to Figure IX.

[68] See Nagel, 1995; Crawford and Iriberri, 2007a,b

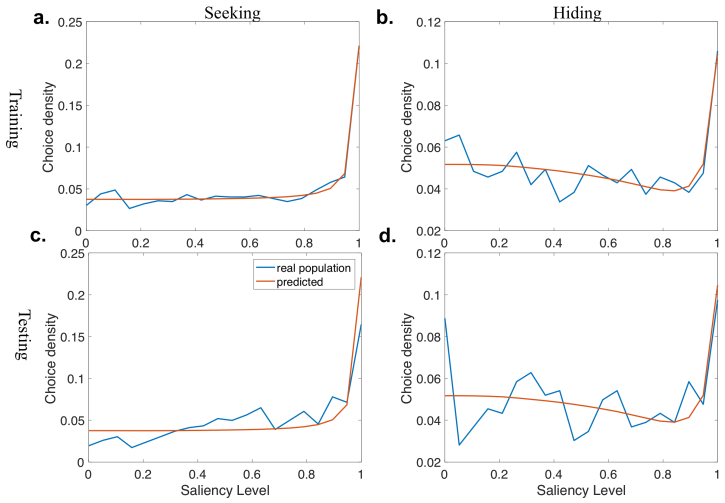


FIGURE E1. LEVEL-K MODEL 6 TRAINING-TESTING COMPARISON

Note: The x-axis is the saliency values of all click points. Each point on a graph indicated what percentage of choices were made for locations within images based on the saliency of those locations. (a): Choice data and model prediction in the training dataset seeking condition. (b): Choice data and model prediction in the training dataset hiding condition. (c): Choice data and model prediction in the testing dataset seeking condition. (d): Choice data and model prediction in the testing dataset hiding condition. Can be compared to Figure IX in the text.

F. WORD LIST VERSION OF MAP COORDINATION TASK

One way to see how important visual saliency is for coordination is to test how people behave when they face the coordination problem in a non-image environment. The SAM algorithm does not apply to such a scenario. We tried a non-visual version of Schelling's location game, in which subjects were asked to coordinate on ten locations described in Figure IIa but only in a word list. The options were: house at the bottom of the map, bridge, small house near the pond, house at the top of the map, pond, two houses together, creek, fork in the road, X on the map, and Y on the map. The questions were presented in a randomized order. N=37 people participated the survey on Prolific. Each of them only answered the question once. Most of them choose the option "x on the map" (49%) while none of them chooses "y on the map" (see table F1). Only 5% people choose the bridge, which was the most popular option when the question

Table F1—The choice percentage of all choices in Figure II

	Percentage
X on the map	0.49
House at the bottom of the map	0.08
Bridge	0.05
Small house near the pond	0.14
House at the top of the map	0.08
Pond	0.03
Two houses together	0.05
Creek	0.05
Fork in the road	0.03
Y on the map	0.00

Note: This table represents the percentage of people (N=37) playing the map game based on a list of verbal description rather than the visual map in text Figure II. Each participants played once.

was presented in an image format.

G. Fruit Experiment

G1. Fruit Experiment -Data

N = 75 participants took part in this study on Prolific, a European online data collection platform, following a pre-registration process on the Open Science Foundation website (OSF). All the participants were pre-screened to have a prior approval rate of at least 70% based on their previous participation. Each subject was only allowed to participate once for all types of batches (including pilot studies). Participation from mobiles and tablets were not allowed in order to control for attention effects.

The experiment design in timeline is shown in Figure G1. Subjects first read instructions freely until they fully understood. They were then asked to answer five comprehension questions as a check. Subjects who made more than one mistake are excluded. Then they played a session with unlimited time to get familiarized with the rules (this part was incentivised also but was only for training purposes, as is not counted in the reported dataset). After that, they entered the main task session, where they would encounter 20 new images in a randomized

order.

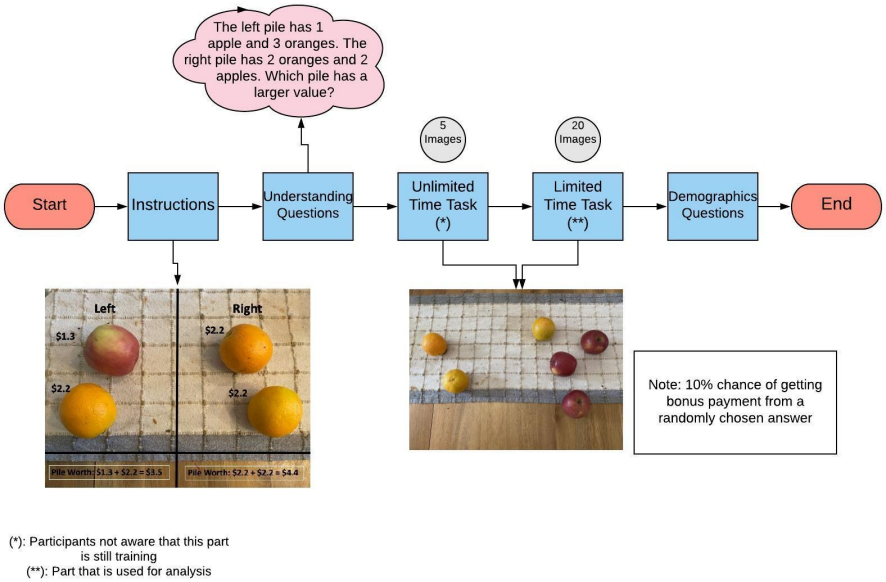


FIGURE G1. EXPERIMENT FLOW.

Note: Subjects first experienced an introduction session explaining the basic tasks followed by a testing session asking questions about the rules. They then experienced an unlimited time session with N=5 images which will count into payment but only for training purposes. Afterwards, they will do a session with time limit. The experiment ends with demographic questions and payment realizations.

G2. Stimuli Properties and Selection Mechanism

We took 72 photos of different combinations of real fruits displayed on a dining table. In the text, Figure III showed examples of SAM predictions. Each image contains two sets of fruits and each set contains three to five fruits. We flipped all the images in the horizontal direction so that we got another 72 images with the same content, but with the set locations flipped horizontally.[69] These 144 images are our image pool.

[69]This procedure is to avoid any left-right biases when taking images. It is done using a matlab function flipimg().

We selected 20 images from the image pool and all of the selected images satisfy four conditions:

1) **One-side salience centered** All of the selected images are strictly one-side salience centered, which means that the most salient locations only appear in one fruit set. Figure IIIc represents an example of a one-side salient image, while Figure IIIa shows an image that is not one-side salience centered. Formally, consider two sets of pixels constituting the left set and the right set, P_l and P_r. Function s, the salience model, maps the union of P_l and P_r to $[0, 1]$. The most salient location of an image consists of a set of pixels $S_h : \{x|s(x) > 0.99\}$.[70] An image is one-side salience centered, if and only if exactly one of the two conditions hold true: $S_h \cap P_l = \emptyset$ or $S_h \cap P_r = \emptyset$.

2) **Balanced salience center locations**: The selected image set has salience centers equally located on the left side or right side. Half of the images have salience centers on the left and the other half have them on the right.

3) **Balanced valuation distribution**: There are only two types of fruits: oranges and apples. Each apple is worth 1.3 dollars and each orange is worth 2.2 dollars.[71] The total value differences between two sets range from 0.4 dollars to 4 dollars. There are exact 50% of rounds with the more rewarding option located on the left and 50% of rounds with the more rewarding option on the right.

4) **Balanced congruences**: An image will be called "congruent" if the more rewarding option is also the more salient option. Among all images, there are 50% congruent images and 50% incongruent images. No image contains two sets with the same amount of values.

[70] Since salience is a relative measure, there will always be at least one pixel with salience value one.
[71] We did a pilot experiment with integer unit values. It turned out to be that integer values were too easy for the subjects so we didn't see any variation in choice accuracy.

5) **Balanced number of fruits**: Among the total 20 images, in 18 images have the number of fruits only differ by one. The other two images differ by two. 11 images have more fruits on the left and 9 images have more fruits on the right.

H. Experimental procedures of location games

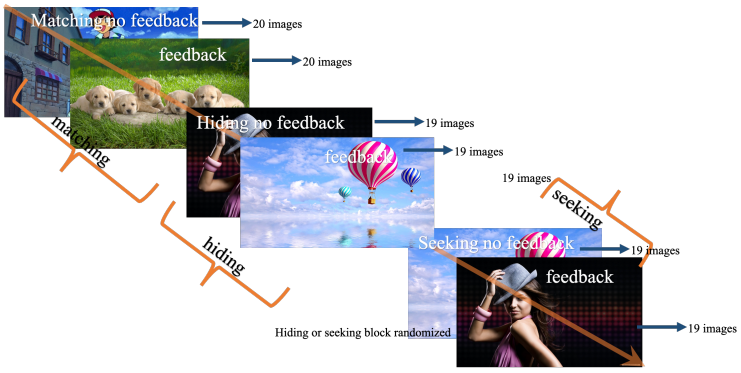


Figure H1. Block Design of The Location GameExperiment

Note: This figure shows the block design of the main experiment (location game). Each participants experienced matching game first, then hiding or seeking game in a randomized order. Under each game, there are two sub-block, the first oen is always without feedback and the second one is with feedback.

Screen1: You are now going to do a series of short games. In each one, you will see a series of pictures and you must choose a location on the picture by clicking with the mouse.

The rules of each game are slightly different, so read them carefully before you start! (You cannot go back and reread them.)

Screen 2: You'll start with a few practice items to help you get familiar with the basic set-up.

Use the mouse to click a location anywhere on picture. Notice that your selection is the entire area within the circle.

You will have 6 seconds to make your selection before the picture disappears. If you do not make a selection within 6 seconds, you will not get credit for that

TABLE H1—SUMMARY OF DATASETS

Type of datasets	**Platform**	**Whether no-feedback**[73]	**N of subjects**	**Games**	**Time limit**	**Between vs Within**
Main Batch	In lab	Yes	29	M,H,S	6 s	Within
Main Batch Online	MTurk	Yes	38	M,H,S	6 s	Within
Big Circle	MTurk	No	67	M,H,S	6 s	Within
High Reward	MTurk	No	29	H,S	6 s	Within
No time-limit	MTurk	No	49	M,H,S	Inf	Within
Between-Subject	MTurk	No	53	H,S	6 s	Between
Time pressure	MTurk	Yes	31	M,H,S	2 s	Within

Note: This table summarized seven different datasets collected at different times. Only the high reward group and the main batch group are the same group of participants. All other batches are completed by a new group of people. Repeated participation is not allowed in all other batches.

picture.

Screen before each session depending on games: Matching: Now you are playing a matching game with several other research participants like you.

For each image, you will play against a randomly selected opponent. If you and your opponent choose the same location in the picture, you both win $x.[72] If there is any intersection between your location and your opponent's location, it will count as a "match".

You won't find out how much you won in this phase until the end of the game.

As before, you will only have 6 seconds to make your choice for each image.

[72]The value of x changes with games, we pay $0.2, $0.1, $0.4 for a success in matching, hiding and seeking.

Table H2—Location games: dataset usage summary

Analysis Names	**Dataset Used**	**Observations**
Seeking Win Rates	The main results: in-lab dataset.	M:559,H:529,S:531
(Seeker's advantage)	Also reported this percentage for other robustness checks.	M:458,H:441,S:452 (Main Batch Online)
SCH model: training	In-lab dataset with both feedback group and no feedback group.	H:1096,S:1090
SCH model: testing	In-lab dataset, high reward group.	H=446,S=446
Choice saliency level analysis	In-lab dataset with both feedback and no feedback group (in footnote and appendix).	M:1139,H:1096,S:1090
Matching rate/Saliency center	In-lab dataset both feedback group and no-feedback group.	M:1139

Note: This table summarizes the dataset we used for each part of the analysis. We mainly and consistently use the dataset we collected in lab for all the analysis. For the seeker's advantage part, we also tested different conditions for robustness checks. The "observation" column denotes the total number of observations under each game. M,S,H denotes for matching, seeking and hiding, separately. We omit all the missing data which happens rarely in the in lab sessions and more commonly in online sessions.

I. Matrix Game

N=56 people played 32 normal form games with different strategic structures: Dominant Solvable Self (DSS), Dominant Solvable Other (DSO), Prisoner's Dilemma (PD), and Stag Hunt (SH).

In the 32 games, 24 games contain a unique equilibrium (SH has two). Each player is either assigned as a row player or a column player. Both roles saw the original games without any transpose.

Table I1—Evaluation of SAM on matrix game experiment

	AUC	CC
SAM vs fixations(games)	0.96	0.47
Chance level	0.5	0
Range	(0,1)	(0,1)

Note:
The table reports two common evaluation metrics for the matrix game experiment in Section VI. It reports area under the receiver operating characteristics(AUC) and Pearson Correlation Coefficient (CC)(Kummerer et al., 2018) . We show SAM's performance on human eye-fixations for matrix games. The results on both metrics show that SAM predicted human fixations far better than chance.

J. Additional analysis

Table J1—Summary of acronyms

SAM	The saliency model we used, Saliency Attentive Model
CI	Crawford and Iriberri (2007)
ABAA	Hider- seeker game using these four letters
CH	Cognitive hierarchy model
SCH	Saliency Cognitive hierarchy model
BGS	Bordalo, Gennaioli and Shleifer (BGS) saliency theory

Note: If readers have difficulty keeping track of all the acronyms, this table may help.