Perception of Ambiguous Drawings and the Construction and Inhibition of its Alternative Interpretation – Reflections on Consciousness

Svetoslav Bliznashki (valsotevs@gmail.com)
Mariya Popova (popova.mariya@gmail.com)
Boicho Kokinov (bkokinov@nbu.com)

Central and East European Center for Cognitive Science, Department of Cognitive Science and Psychology, New Bulgarian University 21, Montevideo Str., Sofia 1618Bulgaria

Abstract

Based on some of our previous experimental findings concerning the negative priming which occasionally occurs during perception of ambiguous drawings we propose a simple mathematical model which we believe may account for important aspects of our data. We tested some (but not all) of the crucial predictions of our model using a newly developed set of asymmetric ambiguous drawings. The results further supported the phenomenon of negative priming taking place during perception of ambiguous drawings and were consistent with the predictions of the model.

Keywords: ambiguous drawings; negative priming; vector reflection.

Introduction

Early gestalt psychologists pointed out that human beings are incapable of perceiving two interpretations of an ambiguous drawing simultaneously. Only one of the interpretations seems to be consciously available at any given moment. Numerous studies show that the conscious experience of a particular interpretation of an ambiguous drawing is subject to context effects (e.g. Balceits & Dunning, 2006; Fisher, 1968; Long & Olszewski, 1999, 2004; Rock & Mitchener, 1992; Kokinov et al., 2007, 2009), i.e. depending on the context we can experience either of the two interpretations. These context effects raise the question whether both interpretations are built in parallel but only the stronger is experienced, or depending on the context only the relevant interpretation is built. This question is analogous to the question asked by Swinney (1979) whether both meanings of a homophone are accessed when the word is encountered or only the meaning relevant to the context. The answer provided by Swinney is that both meanings are accessed. His explanation is in terms of an autonomous lexical access device that does not interact with the context. However, in the case of perception of a picture it is impossible to think in terms of access to pre-stored meanings of the picture, it is clear that the interpretations have to be constructed on the fly. Thus if it turns out that the alternative interpretation is also primed, that would mean that both interpretations are constructed in parallel. Such results would potentially cause a re-interpretation of Swinney’s results since that would mean that this is not a specific linguistic phenomenon that could be explained by the specific organization of the language system and more general explanations should be sought for.

Kokinov, Biznashki, Kosev, and Hristova (2007) were interested in how analogy could cause re-representation and re-interpretation of an ambiguous picture. In a further extension of Experiment 1, Biznashki and Kokinov decided to use additionally Swinney’s methodology (a post-test Lexical Decision Task, called LDT for short) in order to evaluate whether the unseen interpretation will be primed. If this was the case, that would be evidence in favor of the idea that parallel processes are constructing both interpretations. The results of this experiment confirmed the hypothesis of parallel representation building of both interpretations, however, they were rather surprising since we obtained negative priming, i.e. we obtained evidence that the unseen interpretation was inhibited (Kokinov, Vankov, Bliznashki, 2009). This was in accordance with some simulation data with the AMBR model which also exhibited inhibition of the corresponding concepts.

We were happy with these results, but there was a puzzle that remained unresolved and which we found very challenging. In this experiment we used an asymmetric drawing with one interpretation being easy to perceive and the other one being hard to notice. We also varied the context in which the drawing was provided by pushing the participants in the study to perceive either the easy or the hard interpretation in order to solve a complex analogical task. We were happy to prove that the relational structure of the task exerted strong influence upon subjects’ perception (which was the main goal of our studies) but we were puzzled to find out that the negative priming was not always present. More specifically the negative priming occurred if and only if subjects were contextually prompted to perceive the weak (hard-to-see) interpretation of the drawing. In that case the strong (easy-to-see) interpretation showed pronounced negative priming during a Lexical Decision Task which followed immediately after the presentation of the ambiguous drawing, while if the participants saw the easy to see interpretation, there was no negative priming of the alternative interpretation. In our current study we try to explain that finding in formal terms as well as to replicate the results in an experiment specifically designed to detect priming of an unperceived interpretation of ambiguous drawings (which was not the case with our previous studies which focused primarily on exploring the parallel nature of different processes which took place during analogical reasoning).
Mathematical Model and Predictions

Modeling our results was a challenging enterprise. To see why consider the simple competitive learning network (e.g. Knight 1990) shown in figure 1.

Let’s suppose such a network is trained to classify two sets of drawings – one set consisting of drawings of rabbits and another consisting of drawings of dogs. Now suppose that after a successful learning phase the network is presented with a linear combination of two exemplars with the weights of the combination reflecting the degree to which the elements of the input vector retain their similarity to the original exemplars. In most cases the network will “see” the easier interpretation and (given a relevant activation function) will inhibit the harder interpretation. How can we make the network “see” the harder interpretation? One way we might achieve that is by allowing a context node to pour activation towards the relevant output node. This would be analogous to our previous experiments during which we pressed subjects to see the harder interpretation of an ambiguous drawing by placing the drawing in a particular context. As long as the inhibitory connection/s between the output nodes remain constant and uninfluenced by the inputs to the network, however, all our manipulations will result in the more relevant interpretation winning the competition and the inhibition of the alternative. This contradicts with our data since as already explained negative priming seemed to be present only for the strong interpretation when the weak one was supported by context. Because of that we tried to explain our data by a model which will not rely on constraint satisfaction.


In SVRACA we regard concepts as orthogonal vectors residing in a high dimensional space. Each newly perceived entity is represented as a linear combination of the already existing concepts in memory. The vectors representing concepts are considered analogous to Long Term Memory (LTM) while vectors representing newly perceived entities are supposed to be analogous to the content of Working Memory (WM). SVRACA makes a clear distinction between conscious and subliminal perception. Subliminal perception is modeled by simply constructing a vector in WM as a linear combination of concepts/percepts in LTM. Conscious perception is modeled by the categorization of the constructed vector as an example of one of the existing concepts in LTM. The vector from LTM which shows the highest positive correlation with the vector in WM is supposed to represent its designated category. In cases of ambiguity the newly constructed vector in WM exhibits non negligible positive correlation with more than one vectors residing in LTM. In such cases SVRACA reflects the WM vector along all axes (LTM vectors) which the WM vector correlates positively with except for one. The one exception is the single axis which the WM vector still correlates positively with after the reflection operation. All other axes correlate negatively or negligibly with the WM vector. In that state the system is said to have resolved the ambiguity. In other words conscious perception of an entity is said to occur when a WM vector is fully constructed and when it is roughly orthogonal or negatively correlated with all vectors in LTM except for one with which it exhibits a relatively high correlation. Thus in SVRACA vector reflection is considered a metaphor for consciously perceiving only one interpretation of an ambiguous stimulus at any given time. In SVRACA the main operation – reflection is executed in a probabilistic fashion (i.e. a stochastic process decides around which axes the WM vector is reflected and which axis remains constant) and in determining the probabilities of reflection both the internal and external context play crucial role. Internal context (or perceptual context) is represented by the weights of the linear combination which constitutes the WM vector. The larger a weight the more probable the WM vector will not be reflected along the axis the weight is associated with. External context is represented by a vector of probabilities with each probability reflecting the relative strength with which environment pushes towards a particular interpretation. The interaction between internal and external context probabilistically determines which interpretation of an ambiguous stimulus is perceived. Figure 2 illustrates how SVRACA works and some of its predictions:

![Diagram of SVRACA](image)
SVRACA will reflect A either around the X axis (this reflection is indicated as r1 in the figure) or around the Y axis (r2). In the first case the result is A' – a vector unambiguously perceived as X while in the second the result is A'' – a vector unambiguously perceived as Y. Which reflection takes place – r1 or r2 is determined by taking into account both perceptual and external context. Ignoring external context for the moment we see that A is much closer to X and consequently it will be much more probable that r1 occurs. If that is the case A’ will result and X but not Y will be considered to have been consciously perceived. A’ now is practically orthogonal to Y (a very small and possibly undetectable negative correlation exists between A’ and Y). Thus the system has perceived the strong interpretation (X) of the ambiguous stimulus (A) and the alternative interpretation (Y) is neither primed nor suppressed. In the other scenario r2 occurs (because of external context or by pure chance) and A” is perceived unambiguously as Y. Here however the unperceived interpretation (X) correlates very strongly and negatively with A'', i.e. SVRACA predicts that in this case (perceiving the weak interpretation for whatever reason) the unperceived interpretation will be negatively primed. That is exactly what we have observed in our previous studies. It may seem quite arbitrary that we chose vector reflection as the main operation which deals with ambiguity so let’s state our reasons at the very beginning. First of all reflection is a computationally cheap operation. Second it makes a direct use of a meaningful representation of ambiguous stimuli. If an ambiguous stimulus is represented as a linear combination of unambiguous stimuli then reflection is probably the easiest way to attain an unambiguous representation while maintaining the base representational scheme. Third, reflection allows us to model our stimuli as vectors in a high dimensional space which in turn opens the door for exploiting many of the advantages of using highly distributed representations. Fourth, vector reflection may serve as a useful and mathematically sound metaphor for conscious experience. Another such metaphor is reaching a stable state in a recurrent neural network but as we already saw at least some neural architectures don’t provide a straightforward way of explaining the selective negative priming which is the topic of our study. Moreover it is not at all inconceivable representing SVRACA as a complex recurrent neural net in the future. Last but not least a model based on vector reflection not only makes sense of currently available data but also makes useful and testable predictions which will be discussed later.

SVRACA: semi formal description.

Currently SVRACA supports 10 LTM vectors residing in a 10000 dimensional space. LTM vectors are orthogonal to each other and can be said to form a basis of a 10 dimensional space. Each vector is standardized to have a mean of 0 and a standard deviation of 1. Thus all vectors can be conceived as being of unit length. Here we will discuss the simplest case of ambiguity where a WM vector is a linear combination of only two LTM vectors (e.g. an ambiguous drawing which can be interpreted as a rabbit or as a dog). A simulation begins with the system being supplied with a vector of coefficients w which serve as weights for the linear combination. In the simplest case the system is supplied with two coefficients. The sums of squares of the coefficients must be equal to 1 (||w||=1) and thus the length of the linear combination is also equal to 1. The system proceeds by forming a linear combination of as many vectors from LTM as there are coefficients in w. This construction stage reflects encoding a stimulus by the perceptual system and it requires a certain amount of time. Our simulations involve applying w to randomly chosen vectors from LTM since the next stage involves SVRACA determining the relationship of the WM vector to each of the LTM vectors. SVRACA performs a linear regression on WM and collects all coefficients larger than some threshold value (e.g. 0.01). The vectors in LTM corresponding to these coefficients are those which make a significant contribution to the linear combination in WM. The larger a coefficient the closer the WM vector to the LTM vector associated with that coefficient. If all coefficients in w are close to 0 (i.e. below a threshold of 0.1) or negative, SVRACA identifies the WM vector as a previously inexperienced entity and the simulation stops (see the lower left quadrant in figure 2). If all entries in w are close to 0 or negative except for one then SVRACA identifies the WM vector as belonging to the category in LTM associated with the only positive coefficient (see the lower right and the upper left quadrants in figure 2). If more than one entry in w is positive SVRACA interprets the situation as ambiguous and tries to resolve the ambiguity by reflection (see the upper right quadrant in figure 2). In the absence of any external context SVRACA reflects the WM vector around all axes associated with positive coefficients except for one. The probability of any axis remaining unaffected by the reflection operation is proportional to its squared coefficient. For example the probability of WM being reflected around all axes but the third is equal to w(3)^2. External context is represented by a set of weights each ranging from 0 to infinity. The elements of the vector of context weights, denoted c, are subject to only one constraint namely that each of them should be 0 or positive. When external context is present the probability p of each axis being the only one unaffected by reflection is equal to: 

\[ p = \frac{((w.*c+w).^2)/\text{sum}(w.*c+w).^2)}{1} \]  

(1)

where “.*” designates element-wise operations, p is the vector of probabilities that each axis is the only one unaffected by reflection, w is the vector containing the weights determining the WM vector and c is the vector containing the weights representing the strength of external context towards each LTM axis. After a single axis is chosen to be unaffected by reflection SVRACA reflects the WM vector around all other axis with positive weights. This is achieved by setting to 0 all elements in w which were initially negative as well as the element corresponding to the unaffected axis. This newly formed vector is called...
reflection vector and is denoted by wr. Now all positive elements in wr refer to axes WM is to be reflected around and all other elements are 0. If we denote the collection of vectors in LTM which are involved in the linear combination WM as V (V is now a matrix with dimensions 10000xlength(w)) the reflection WMt is achieved by:

\[ WMt = WM + V^*(-2*wr) \]  

(2)

where WMt is the reflected version of WM, V is the collection of all vectors in LTM originally participating in the formation of WM and wr is the reflection vector. The final step in a SVRACA simulation is to examine the relationship of WMt to the vectors in LTM in order to decide which LTM concept is the only positively related one to WMt (i.e. which concept WMt is perceived as). The relationship between any two concepts in SVRACA (LTM, WM, WMt) is defined as the cosine of the angle between the vectors in question. Since all vectors are defined as being of unit length the cosine of the angle between two vectors is also equal to the product moment correlation between them. Thus after reflection WMt will in most cases be positively correlated with only a single vector in LTM, negatively correlated with all other vectors in LTM which participated in the formation of WM and orthogonal to all vectors in LTM which were not involved in the construction of WM in the first place. The signs and magnitudes of these correlation coefficients are supposed to predict priming (positive or negative) in real world situations.

A Typical Simulation is described next. We supply SVRACA with a highly asymmetrical concept which is a linear combination of two LTM vectors with coefficients \( w = [0.95 \ 0.3122] \). We see that the WM vector is much closer to the LTM1 vector than to LTM2 vector. The context vector, however reflects the opposite picture: \( c = [0.5] \), i.e., the LTM1 vector is absolutely irrelevant to the situation while the LTM2 would be very useful. Since both numbers in w are positive SVRACA recognizes the ambiguity (i.e. identifies the LTM vectors correlating positively with WM) and calculates the probabilities p of consciously perceiving each of the two possible interpretations according to (1): \( p = [0.2045 \ 0.7955] \). We see that the contextual influence radically changes the odds towards a conscious perception of LTM2. Next SVRACA stochastically determines around which axis to reflect WM based on the just obtained probabilities. Among 100 simulations with the same input parameters SVRACA interpreted the ambiguous WM as LTM1 16 times, and as LTM2 84 times. The reflection operation (2) resulted in WMt correlating -0.3122 with LTM2 and 0.95 with LTM1 in the 16 cases (conscious perception of LTM1) and 0.3122 with LTM2 and -0.95 with LTM1 in the 84 cases (conscious perception of LTM2). If we take the square of these correlation coefficients while keeping in mind their signs we see that when LTM1 is consciously perceived LTM2 remains practically unprimed – its overlap with WMt is equal to -0.0975. In the second case however when LTM2 is consciously perceived LTM1 exhibits pronounced negative priming – the magnitude of the negative overlap (i.e. the signed coefficient of determination) between WMt and LTM1 is -0.9025. In other words SVRACA predicts practically no priming of the unperceived interpretation when the system perceives the strong interpretation and strong negative priming of the unperceived interpretation when the weak interpretation is perceived due to external context. We proceed with an empirical study testing these predictions.

Experiment

Design and Stimuli

We attempted to test our current predictions and further verify our previous findings by developing a new set of ambiguous drawings. We developed 3 pairs of asymmetrical drawings. Each pair contained an easy-to-see and a hard-to-see version of a particular picture. The first pair of ambiguous drawings can be seen as either a rabbit or a dog (figure 3). The two variants of this drawing included an easy-to-see rabbit (hence a hard-to-see dog) and an easy-to-see dog (hence a hard-to-see rabbit). The second ambiguous drawing can be seen as either a duck or a goat. The third drawing embodied a mouse and a frog and its two variants depicted an easy-to-see mouse (a hard-to-see frog) and an easy-to-see frog (a hard-to-see mouse).

Figure 3. The two variants of the first ambiguous drawing.

Table 1: Percentages of people who saw each interpretation

<table>
<thead>
<tr>
<th>Version:</th>
<th>% who saw</th>
<th>% who saw</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strong rabbit</td>
<td>95</td>
<td>5</td>
</tr>
<tr>
<td>(weak dog)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strong dog</td>
<td>78</td>
<td>22</td>
</tr>
<tr>
<td>(weak rabbit)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strong duck</td>
<td>71</td>
<td>29</td>
</tr>
<tr>
<td>(weak goat)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strong goat</td>
<td>95</td>
<td>5</td>
</tr>
<tr>
<td>(weak duck)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strong mouse</td>
<td>84</td>
<td>16</td>
</tr>
<tr>
<td>(weak mouse)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The variants of each drawing were validated in a simple picture naming task in which each variant of each drawing was presented on a computer screen (among many fillers
depicting unambiguous drawings) and subjects were simply required to name what they saw. Each subject saw only one version of an ambiguous drawing (to avoid priming effects) and each version was seen by 100 subjects. Thus 200 subjects participated in the standardization procedure. Table 1 shows the percentages of people who interpreted each version in each way.

For each variant of a picture we compared statistically (chi squared tests) the percentages of people seeing the strong and the weak interpretation of a drawing and all comparisons yielded highly significant results. Thus our stimuli indeed represented asymmetrical ambiguity. In the study we used the drawings as follows: each subject was presented with a single version of each ambiguous drawing only. The drawing was placed in a context which favored either the easy or the hard interpretation. There were three conditions in a repeated measures design: in the first condition subjects were contextually prompted to see the easy version of a drawing; in the second condition subjects were prompted to perceive the hard version; in the control condition subjects saw an unambiguous version of a drawing.

Procedure
Each subject saw three drawings in three conditions in three target trials. Each target drawing appeared as a possible solution to a simple analogical task of the type A:B::C:? with two possible answers. The ambiguous drawing was one of the possible answers and the only way to solve the task correctly was to perceive the ambiguous drawing in its intended interpretation. The analogical task served as a context which prompted participants to perceive consciously only one of the possible interpretations in each drawing. In the first condition (called Easy Condition or EC from now on) subjects were required to perceive the easy interpretation of a drawing in order to solve the task correctly. In the second condition (called Hard Condition or HC) subjects were required to perceive consciously only the hard interpretation of a drawing. In the third condition (Control Condition or CC) subjects were presented with an unambiguous version of the drawing which was obviously the correct response to the analogical task. Each subjects saw only one version of an ambiguous drawing in order to avoid priming effects. Subjects were required to respond with the solution to each analogical task verbally in front of a microphone by naming the option they felt best fitted the analogy. After the response the microphone triggered the start of a LDT. A word appeared at the center of the computer screen and subjects were required to indicate whether it was a meaningful word in Bulgarian or a random sequence of letters by pressing a button. After each target trial a meaningful word appeared which referred to the supposedly unperceived interpretation of the just presented ambiguous drawing. In the CC where an unambiguous version of one of the ambiguous drawings was presented the word in the LDT referred to the other interpretation of the corresponding ambiguous drawing. The timing of the study was as follows: an analogical task remained on the screen until the subject responded to it in front of the microphone. Immediately after the response a white screen with a fixation cross at the center appeared and remained for 350ms. After that a string of letters appeared at the center of the screen and remained there until subjects indicated by pressing one of two buttons whether the letters comprised a meaningful word or not. After the response a blank screen lasted 7sec. before the beginning of the next trial. The experiment contained three target trials and seventeen filler trials (the filler trials were the same as the target ones with the exception that the analogical tasks didn’t contain ambiguous drawings and the LDT following them could contain both words or non-words after each analogical task; in contrast only words followed the target trials). The order of presentation of the three conditions as well as the order of presentation of the three drawings and the variants of each drawing were counterbalanced between participants resulting in 24 different lists. Each subject participated in only one list. We give examples of the first two lists in order to clarify the procedure: A subject in the first list was firstly supposed to see an easy duck (i.e. the easy-to-see duck variant of the duck/goat drawing was presented in a task that required the subject to see a duck in order to solve the task correctly) – EC, than she was supposed to see a hard frog (i.e. a hard-to-see frog variant of the frog/mouse drawing was presented in a task that required the subject to see a frog in order to solve the task correctly) – HC and finally she saw an unambiguous dog in a task that required her to see a dog in order to solve the task correctly – CC. In the EC condition she responded to the word “goat” during the LDT, in the HC she responded to the word “rabbit”, and in the CC she responded to the word “mouse”. A participant in the second list was firstly supposed to see a hard dog (respond to the word “rabbit” during LDT) – HC, then she was supposed to perceive an unambiguous goat (respond to “duck”) – CC and finally she was supposed to see an easy frog (respond to “mouse”) – EC. The remaining 22 lists exhausted all other possible combinations of variants of pictures to be perceived (and hence words to react to during LDT) and orders of presentation of conditions. In total we collected Reaction Times (RT) to six words all appearing in each of the three conditions. Those RTs constituted our dependent measure.

Participants
Sixty eight undergraduate students (45 females and 23 males) from New Bulgarian University participated in the study for obtaining credits.

Results and Discussion
Fifteen subjects failed to perceive the hard-to-see interpretation in the HC condition and were replaced. Five subjects were replaced because some of their RTs to target words during the LDT exceeded our threshold of 1500ms. After replacement of those subjects we were left with 48 participants – exactly two participants per list. The decision
to have two valid participants per list was made prior to the beginning of the study. The average RTs for each condition are presented in table 2.

Table 2: Average RTs for the LDT in each condition.
Standard deviations are presented in parentheses.

<table>
<thead>
<tr>
<th>Condition</th>
<th>RT</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control Condition (CC)</td>
<td>774ms</td>
<td>134</td>
</tr>
<tr>
<td>Easy Condition (EC)</td>
<td>777ms</td>
<td>130</td>
</tr>
<tr>
<td>Hard Condition (HC)</td>
<td>824ms</td>
<td>146</td>
</tr>
</tbody>
</table>

We analyzed our data with a linear mixed model (e.g. Hoffman, 2007; Brysbaert, 2007) in which we entered subjects and items (i.e., words in the LDT) as random factors simultaneously. Our design allowed only for the inclusion of random intercepts for each random factor. The fixed factor was “Condition” with three levels for each subject (EC, HC, CC). The analysis showed a highly significant overall result for our independent variable – F(2, 89.301)=5.718, p=0.005. Multiple comparisons (performed via the SIDAK method) revealed a significant difference between CC and HC (p=0.01), EC and HC (p=0.017) and virtually no difference between CC and EC (p=0.997). We see that the results corroborate our previous findings and are perfectly consistent with the predictions of SVRACA: when subjects perceived consciously the hard-to-see interpretation words referring to the easy-to-see interpretation become negatively primed compared to a control condition involving no ambiguity but not vice versa (i.e. no priming is detected when subjects perceive consciously only the easy-to-see interpretations). It is important however not to overgeneralize these findings. Fifteen subjects in our study (22%) failed to perceive the hard-to-see interpretation in the HC (i.e. perceived only the easy interpretation or reported perceiving both interpretations) even in the presence of strong context. This wasn’t unexpected since our study involved a condition in which subjects were required to perceive consciously only a very hard interpretation of an ambiguous figure. No contextual influence exists which can assure that this happens 100% of the time. Nevertheless we feel compelled to point out that given these circumstances our findings extend only to subjects which were “context sensitive” enough to comply with our experimental manipulation. The future directions of our work include testing other explicit predictions of SVRACA. Such predictions include: priming positively only the strong interpretation of an ambiguous drawing when the figure is presented subliminally; smaller effects compared to the just presented when symmetric ambiguous drawings are involved (i.e. when dealing with drawings with two equally easy to perceive interpretations SVRACA predicts that negative priming will again occur due to reflection but the amount of priming will be considerably smaller since reflection of a bisecting angle vector will result in a smaller negative correlation with the unperceived concept in LTM); generally smaller effects of positive priming during subliminal presentation of ambiguous stimuli.

The obtained results are coherent with analogous results obtained by Nievas and Mari-Beffa (2002) who discovered negative priming of the non-selected meaning of a homograph, but only if the non-selected meaning is the dominating one. The combination of their and our data shows that the phenomenon of negative priming of the alternative representation is broader and holds both for linguistic and perceptual task and thus should be interpreted as unconsciously building a representation of both alternatives and inhibiting the stronger one if for contextual reasons we are pressed to use the weaker interpretation.

References