The Effects of Meaning in a White/Black Setting on Reasoning

Eshaa M. Alkhalifa
Department of Computer Science,
College of Information Technology, University of Bahrain,
P. O. Box 32038, Isa Town, Bahrain
ealkhalifa@sci.uob.bh

Abstract

The Wason (1966) selection task has persistently resulted in weak subject performance if the question is abstract and does not imply a 'role' must be performed. Some work (Alkhalifa, 2001b) has shown the effect of 'directing' subject behavior towards selecting p or notq using a conveyor belt system where each experiment was only able to produce results in one direction. Here, the same conveyor experiment is repeated after changing the colors of the objects to black and white resulting in a balance in selections between p and notp. This balance, is caused by color, so it is an effect of 'meaning' and it over-rides the effect of directionality implying that the latter too was an effect of meaning. These results offer an explanation to the mystery in the abstract form of this task, by showing that what influences subject performance in the abstract task is 'meaning' represented through direction. Yet this effect can be neutralized if another semantic effect exists caused by opposite colors that interfere with the implied directionality. These two colors seem to imply a sense of balance as with two weights on a scale denying the possibility of any fluctuation in the weights and restricting the implication of direction.

Introduction

The card task proposed by Wason (1966) was composed of four cards presented to students who were told that the cards had a number on one side and a letter on the other. They would then be told to check the truth or falsity of a rule as in If A then 7 and given the cards showing an "A", a "B", a "4" and a "7". They would be asked to turn over or select the cards they wish to turn over to check the rule. The correct responses would thus be A and 4. However, only 4% to just over 20% of students select these two cards (Manktelow & Evans, 1979).

The questions raised, resulted in three main streams of work. The first group seems to question the accuracy of human logic and its proneness to error. The second group investigates the mental representation that may be responsible for the types of responses students give. The third group seems to be more interested in investigating what types of materials students would perform better with and what can be inferred from the results. Three streams that pour into the same beautiful river as every effort or attempt is worthy of consideration.

The first group assumes a logically based mental representation where any behavior that does not conform to what is expected by logic is considered an 'error'. Abstract rule theories of mental representation fall under this category by assuming that abstract rules are utilized for logical inference. Braine's theory (1978) as an example, gives three possible reasons for logical errors including comprehension errors, heuristic inadequacy errors and processing errors.

The second group, on the other hand, accepts these errors due to their high frequency as a natural consequence to the representational system and how processes information. The Mental Models theory (Johnson-Laird & Byrne, 1991, 1993) attempts to make such an explanation by proposing a representational system that generates mental 'models' to represent the problem at hand, based upon the meanings of the premises and general knowledge. The claim is that it is based on the Principle of Truth.

"A fundamental assumption of the theory is the principle of truth: Individuals minimize the load on working memory by tending to construct mental models that represent explicitly only what is true, and not what is false" (Johnson-Laird, 1999)

The third group attempts to study the effects of the various materials that can be used in this format without necessarily attempting to explain the results within a particular framework. Work done under this umbrella, is dispersed and subject to a diverse medley of interpretations that free biases that force them to conform to any particular theory of representation. However, this opens the door to a variety of starting points that cover a wide spectrum the most prominent of which at the moment is the sensitivity of the task to Bayesian considerations. However, efforts do exist to make the Mental Models theory, the representational theory to account for the results explained through probabilities (Johnson-Laird, 1999).

Almost all the work done in the above three streams, pour into one extremely important conclusion, namely the effect of "checking for cheaters" or giving subjects an idea that the rule might be broken.

"Selections are sensitive to the probability of encountering a falsifying instance." (Johnson-Laird, 1999)

This led to a division of the materials into two main groupings, thematic or content materials that may carry some sort of implication that a rule may be broken and abstract materials that do not. 'Thematic' materials were observed to result in a much higher level of accuracy (Wason & Shapiro 1971, Wason & Johnson-Laird 1972). These materials were characterized with a thematic relationship between the two propositions in the conditional. An example of this type of problem is: "If I go to Manchester, I go by train" (Wason & Shapiro, 1971). On the other hand, 'abstract' materials included the ones presented in the classical Wason selection task usually characterized by being 'concrete' but lacking in any thematic linkage between the propositions. However, the dividing line was never clear-cut and almost no insight was gained to illuminate the possible causes of the poor performance in the abstract or arbitrary tasks.

"Some of these patterns are inconsistent with any logical interpretation of the materials."(Cheng & Holyoak, 1985)

Experimentation went on to define the types of thematic linkage that would elicit more accurate responses leading to the effects of having a 'rationale'. Cheng and Holyoak, (1985) proposed that people use 'pragmatic reasoning schemas' and when they are given a 'goal' to seek they would perform better in thematic tasks. They also attempted an "abstract" permission schema where subjects were told that they were an authority checking whether or not people obey a certain regulation (Cheng and Holyoak, 1985). Another similar competing theory is that of 'social contracts' introduced by Cosmides (1989). The contract relates perceived benefits to perceived costs. In a sense, it works as an exchange of cost for benefit and cheating occurs when one fails to pay the cost and accepts the benefit. Gigerenzer and Hug (1992) further supported this theory by showing that when subjects are asked to adopt a 'social role perspective' they would perform much better at the same task. One of the rules they tried was "if someone stays overnight in the cabin, then that person must bring along a bundle of wood from the valley." This rule is familiar to mountain hikers and there are rumors that the rule is not always followed. Subjects are cued into the perspective of a guard who checks whether any one of the four hikers has violated the rule by turning cards. Gigerenzer and Hug (1992) claim that all 'social contracts' are either permission or obli-

gation rules but not vice versa. They add that the pragmatic reasoning schemas does not utilize the notions of cheating as an option and that being cheated is relative to the perspective of one party only. Clearly research into thematic materials has shown a clear requirement to perceive both perspectives or at least to entertain the possibility of cheating. This is another way of saying that P may be true while Q is not in order to urge subjects to test the NotQ condition, which they now so often neglect.

Cheng et al, (1986) did an interesting experiment that involved training subjects through 'contingency' versus 'obligation' training. They cases they recorded the worst improvement with percentage errors of 85%, 60% and 58% in the control, contingency and obligation training. The two rules that had this low improvement rate were "If a card has an 'A' on one side, then it has a '4' on the other side" and "If a bird on this island has a purple spot underneath each wing, then it builds nests on the ground". Clearly both rules, do not exhibit any semantic links between the two premises other than spatial location. Yet it is this particular sort of task that seems to exhibit a high selection rate of P with the rule IF P then Q.

An implicit emphasis seems to be consistently made on the importance of directionality in this simple task. Subjects seem to select P with a much higher probability than when they select NotQ. Cheng and Holyoak (1985) also showed that people are able to reason more accurately when the rules are stated in if-then form rather than only-if form.

Directional Thought

Freyd and Finke (1984) proposed a theory of representational momentum. They showed through a series of experiments that when subjects are exposed to a series of still images and later asked to select the final location, they would add a step forward in the direction of movement without noticing it. This behavior seems to indicate that they get 'tuned' in the perspective of a changing scene and then find it difficult to halt that scene at a particular point.

"One of Wason's strongest empirical claims from these studies of insight is that subjects who start with the p and q choices, never do reach a state of complete insight."(Stenning and Lambalgen, 1999)

We find some similarity between the two in that both show a uni-directional behavior pattern that may lead to the selection of P and Q as is the case. Stenning and Lambalgen (1999) emphasize the point that any comprehensive theory that attempts to explain behavior in the selection task must not only be concerned with 'successful cases' but also be concerned with the causes of 'errors'. Therefore, the factors that affect and guide behavior towards making the errors are equally important to the factors that guide behavior towards more correct responses. Following from that we may assume that the simple division of materials into 'thematic' and 'abstract' is not sufficient to explain behavior. Not only that but it seems to indicate two classes of behavior patterns that intersect only when a possibility of a false instance of the rule is shown. The representational consequences of the existence of this falsifying instance remains to this date unexplained.

The two types of materials continue to exist on two islands that are to this date fairly isolated from each other and patiently await the person who will dare build the first bridge between them. But in order to build a bridge, all must pass, abstract as well as thematic materials must be explained by the same bridge. Another way of looking at this bridge is a link that may exist between the two and is shared by the two. A form of transportation that would carry the effect of one type to the other which to this moment seems non-existent, or is it?

Cheng and Holyoak (1985) showed the possible existence of such an effect when they interchanged the order of an 'abstract' permission rule with the abstract card question. They noticed that the card problem was solved more correctly 39% when it followed the permission problem, than when it followed the abstract problem 48%. An analytical study (Alkhalifa 2001a) showed the existence of two main variables

that affect students responses in both types of tasks. They are egoism and temporal implication and are positively correlated to performance in no less than 67 experimental runs that were conducted by well-known scientists. Further evidence was provided (Alkhalifa, 2001b) through the contrast of two conveyor belt settings with objects that are either gray or striped on the conveyor belts. If students are not told that the conveyor is not moving, then the main effect is noticed in the notq condition. Students seem to refuse to entertain the possibility of a notq condition if the conveyor comes first, and entertain it more in the conveyor question if the abstract question comes first. Notice that a form of analogical mapping occurred here, since an 'inference' is carried over from one problem to the next.

By converse, if the students are told that the conveyor is moving, the main effect is detected in P, as selections of P are extremely high in both version but only in the conveyor case. Selections of P are significantly decreased if the abstract question comes first and selections of notP are significantly increased. These findings are regarded as significant since to this date it has been extremely difficult to find an effect that would alter the classical card task behavior. This directional bias, can be achieved merely through presenting a conveyor belt that is either described as moving or static. It may seem trivial to assume that this directional bias is semantic, yet, this can be challenged, if we remember that the two settings were analogically mapped onto each other. Although selections of p not p q and not q were ordered differently in the two questions, the structure of the two questions mapped perfectly. So a seriously important question here is: Is this directionality a semantic trait associated with the conveyor, or is it a trait that conveyors can move and therefore they have direction? In other words, is the directionality of thought, a semantic trait, or a structural one?

The Semantics of the Stroop Effect

The Stroop effect, represents the interface that result when people are asked to name the color in which words are written, when the words are names of colors incongruent with the colors they are written in (Stroop, 1935). The work done to study this effect is phenomenal, and the effect solid. The most relevant finding to the work presented here is that made by Klein (1964) who demonstrated that this interference varies on a continuum as a function of semantic association between relevant and non-relevant dimensions. Klein varied the strength of the semantic association between words and the concept of color, as with words that are strongly associated with color such as "sky" and words that are not associated with color such as "house". MacLeod (1991) indicates that the finding of a semantic gradient of interference indicates that a semantic representation of the word is activated. He also, indicates that the degree of association between this semantic representation and the concept of color modulates the amount of interference.

It should be then safe to conclude that the Stroop effect is a semantic effect that is caused by an incongruence of color, with the semantics of the word. Therefore, if the directional effect that was detected in the conveyor experiments (Alkhalifa, 2001b) is semantic then it should be affected by the Stroop effect. If on the other hand, the directional effect is not semantic, it should resist the effect of altering the color of the objects.

However, a problem arises with respect on how to induce the Stroop effect onto the conveyor task because the task involved has P and notP, Q and notQ conditions rather than having color meanings implied. One possibility would be the choice of a white/black scenario because white is regarded as the negative of black and black regarded as the negative of white.

Experiment

The experiment presented here is aimed at testing the effect of having black/white objects to see if they are capable of affecting the behavior of subjects in the same tasks presented in the conveyor and card

experiment. The results of this experiment would then be contrasted with the results previously found, to identify any effect of the new color setting, when contrasted with the results of the gray/striped material presented in the previous work (Alkhalifa, 2001b). It is predicted that the black/white setting would strongly affect subject performance in maintaining a balance between the choices of P and notP, and between Q and notQ in both the conveyor and card settings. This interference should be sufficient to conclude that the directional effect previously found, is indeed a semantic one.

Design Details

The task was a paper and pen task, presented as a small booklet starting with an introduction page and instructions. Students were instructed to go through the two questions in order, one page at a time, and not to look back.

Subjects

55 high school students from the Indian school volunteered to perform this task. 28 students did the conveyor question first followed by the card question. 27 students did the card question first followed by the conveyor question.

Materials

Two questions were given in alternative order to each of the groups. The abstract question is the classical abstract question formatted in a fashion similar to the second question and the colors of the letters and numbers are altered to a black A white A, black B and a white B. The question becomes: Below are four cards with black letters on one side and white letters on the other side. Your task is to decide which of the cards you need to turn in order to find out whether or not the rule is being followed. The rule is: If a card has a black 'A' on one side then a white 'A' must be on the other side. Turn over only those cards that you need to check to be sure.
The second question is: "Following are four figures that show four conveyor belt systems. Each transports two objects one at a time. A white object first exists from within the conveyor box, appears on the belt and then falls into the loading box to be stored. The white object is followed by a black object that is either on the conveyor belt or still waiting in the conveyor box. You are asked to check the following rule by selecting any of the following cases that need inspection. Inspection involves opening the appropriate box and verifying that it contains the appropriate object. The rule is: If the white cube is in the conveyor belt system, then the black cylinder must be the black object directly following it as the conveyor belt moves. Each of the figures has a sentence as in "Look inside the loading box." Or "Wait for the next object to appear."

Results and Discussion

First.Q	P	Not P	Q	Not Q
A.Q.F	59.3%	59.3%	25.9%	29.6%
C.Q.F	53.6%	67.9%	17.9%	35.7%
P<	.245	.080	.068	.181

Table 1: Abstract Question selections in both conditions

First.Q	P	Not P	Q	Not Q
A.Q.F	37%	25.9%	29.6%	33.3%

C.Q.F	42.9%	42.9%	35.7%	25%
P<	.241	.0006	.202	.055

Table 2: Conveyor Belt Selections in both conditions.

The results shown in the tables should not at all be surprising. Look at the similarities between the P and NotP conditions when the questions of each type appear first. An example would be if the Abstract question appears first, the percentage selection of P is 59.3% and the percentage selection of NotP is equal. This is clearly caused by the color effect of the strict contrast between the black and white colors. The same effect appears when the conveyor task appears first.

Although the two tasks do attempt to affect each other as they did in the previous cited work (Alkhalifa, 2001b), the semantic interference of the colors seems to over-ride most of the effect. The only one that remains strong is NotP and that effect in the previous work when subjected to a chitest resulted in a $p<.000$.

Comparative Study

The results obtained above, support the prediction that an interference of color seems to affect the choices made in the selection task. The opposite colors of black and white seem to result in a balancing effect between the choices of P and NotP and between Q and NotQ. However, it does not give enough of an indication when analyzed alone, on how different the black/white setting is from the striped/gray setting. Note that the conveyor question used in the comparison is the static conveyor question.

Results and Discussion:

Type	P	Not P	Q	Not Q
A.Q.F	85.7%	14.3%	76.2.7%	33.3%
B/W	59.3%	59.3%	25.9%	29.6%
P<	4.4E-14	7.2E-38[1]	3.5E-32	.428

Table 3: Abstract Question selections with both Question types

First.Q	P	Not P	Q	Not Q
C.Q.F	55	65%	15%	15%
B/W	42.9%	42.9%	35.7%	25%
P<	.015	3.6E-6	6.7E-9	.005

Table 4: Conveyor Belt Selections with both question types.

The results compare the experiment conducted with black and white objects to the one that has gray and striped objects. Other than that, the same questions were used, yet the behavior was extremely different simply because of the change of color. Any remaining doubts can be eliminated by looking at what the

---

[1] Significance in the three values varies from $8 *10^{-7}$ to $4*10^{-43}$

black/white setting did to the classical Wason card when compare to the normal black letter on white background setting. Selections of P and NotP are identical while in the classical setting, there is a much higher percentage of selecting P.

Analysis:

- Although the Black/White question is formulated as a Wason selection task question, it failed in resulting in high selections of P and Q that usually exceed selections of NotP and NotQ. The only difference between the neutral task in the previous work (Alkhalifa, 2001b) and this one is color, so the difference must be caused by color.
- The effect of one question on the other that was observed in the neutral case in the previous work is not evident here, so there was an interference with this caused by color.
- In the Stroop effect (1935), the color of the word interferes with the word meaning, with latter seemingly stronger than the former. The Stroop effect is believed to be semantic because a semantic gradient of interfaces was obtained (MacLeod, 1991).
- The color effect observed here also interferes with the reactions of subjects in the task, so do we dare presume it to be semantic by analogy?
- The percentage selections of P and NotP were equal in the abstract task and equal in the conveyor task, yet the two numbers were not equal, so is there a gradient of sorts here as well.
- If P= White Cube and Q= Black Cylinder
Then either:
NotP= Black Cube and NotQ= White Cylinder
Or
     NotP= White Cylinder and NotQ= Black Cube
The difference is in how the negation is applied

- If P is White, NotP is Black, then the effect of color is to equate P and NotP percentage selections. However, if P is White and Q is Black, then the effect of color would be to equate P and Q selections. Since P and NotP are equated then the assumption of subjects could be:
P=White Cube and Q=Black Cylinder
Leading to
NotP=Black Cube and NotQ= White Cylinder

- In the abstract task P selections equal NotP selections and Q selections equal NotQ selections. In the Conveyor question, P selections equal NotP selections but Q selections are close to NotQ but still differ with 10%. Oddly enough, in the neutral Conveyor question has equal Q and NotQ.
- An interesting question here is; why are Q and NotQ selections equal in the neutral Conveyor question and not equal in the Black/White question? There must have been something in the Conveyor question's semantic implications that interfered with allowing the colors to equate the numbers, or was it that it interfered with allowed the conveyor question from equating the numbers?
- They both seek to equate the numbers but they do not seem to be influencing the percentages in the same direction or with the same force, so the results are different. Perhaps one of the two effects presented a resistance to the other.
- If the effects of color on reasoning is considered semantic, yet they are a structural trait of the objects under consideration, then are colors the 'meaning' in structure? Is directionality implied by the structure of the objects presented or described?

General Discussion

This experiment seems to wish to tell more than what is implied at first instance. The effect of altering the colors to black and white seems to imply a 'semantic' understanding of opposites. This 'semantic' information was not self evident when letters were used versus numbers, nor was it self evident when striped objects were used versus gray objects. This information contained in the color of the object, seems to interfere with the directional bias that the conveyor was noted to influence subjects with in the prior publish work (Alkhalifa, 2001b). From the extensive work on the Stroop effect we find that it is defined, as when the meaning of the word interferes with its color. The Stroop effect is identified as a semantic effect because a semantic gradient of interference exists (MacLeod, 1991). If the Stroop effect is semantic and one of the causes of the effect is the color of the word, then isn't it likely that the effect of the white/black color influence is also semantic. After all, they are all colors.

Once, we accept that the effect detected here is a semantic effect of color, we only have to try to relate this to reasoning and the effect of direction. The effect of these colors affected the biasing ability of the conveyor system by interfering strongly with it. It did not allow an otherwise strong effect to be observed. In order to interfere with this effect, then there must be a comparative basis on which the two interact. Isn't it then logical to assume that the two effects are semantic, and interfere with each other if they are not aligned in the same direction? After all, the questions compared are identical except for the Black/White setting versus the Striped/Grey setting. The only factor that was different between the two was the color.

Note that, the Black/White setting had results that are extremely different from all the classical Wason (1966) selection task results. In a sense, alone it showed itself as unconventional. Usually, selections of P are the highest of all, closely followed by selections of Q and the selections of NotP and NotQ are fewer than the above with selections of NotQ usually being the least of all. This behavior wasn't followed in the Black/White question clearly affected by the assumption that the color of Black and White are opposites of each other likes the opposite sides of a coin.

Since this effect interfered with the effect of direction, then this effect too has a direction that is opposite to the other. If the other directed people's thought from P to Q then this, was uniformly affecting the two to balance selections between P and NotP, etc. The results reflect a world of forces that seem to exist, that interact with each other in a complex fashion and resist each other's pressures in an interesting way.

Acknowledgments

References

Alkhalifa, E. M. (2001a), Egoistic Reasoning 'In Time' in the Selection Task, Proceedings of the Third International Conference on Cognitive Science, ICCS 2001, pp. 109-113, Beijing, China.

Alkhalifa, E. M. (2001b), Directional Thought Effect in the Selection task, Proceedings of the Third International Conference on Cognitive Science, ICCS 2001, pp. 171-176, Beijing, China.

Braine, M. D. S., (1978), On the Relationship between the natural logic of reasoning and standard logic. Psychological Review, 85, 1-21.

Cheng, P. W. & Holyoak, K. J. & Nisbett, R. E. & Oliver, L.M. (1986), Pragmatic versus syntactic approaches to training deductive reasoning, Cognitive Psychology, 18, 293-328.

Cheng. P. W. & Holyoak, K. J. (1985), Pragmatic Reasoning Schemas, Cognitive Psychology, 17, 391-416.

Cosmides, L. (1989), The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. Cognition, 31: 187-276.

Freyd, J. & Finke, R. A. (1984) Representational Momentum, , Journal of Experimental Psychology: Learning, Memory and Cognition, 10, 126-132.

Gigerenzer, G. & Hug, K. (1992), Domain-specific reasoning: social contracts, cheating and perspective change. Cognition, 43: 127-171.

Johnson-Laird, P. N. & Byrne, R. (1991), Deduction, Hove, UK : Lawrence Erlbaum Associates Ltd.

Johnson-Laird, P. N. & Byrne, R. (1993), Models and deductive rationality. In K.I. Manktelow & D.E. Over (Eds.), Rationality: Psychological and philosophical perspectives, London : Routledge.

Johnson-Laird, P. N. (1999), Deductive Reasoning. Annual Reviews of Psychology, Highwire Press, Stanford University.

Macleod, C. M. (1991) Half a century of research on the Stroop-like interference in word-naming tasks. Journal of Experimental Psychology: Human Perception and Performance, 11, pp.62-80.

Manktelow, K. I. & Evans, J. St. B. T. (1979), Facilitation of reasoning by realism: effect or non-effect? British Journal of Psychology, 71: 227-231.

Stenning, K. & Lambalgen, M. van (1999), Is psychology hard or impossible? Reflections on the conditional, In J. Gerbrandy, M. Marx, M. de Rijke, Y. Venema (Eds), Liber Amicorum for Johan van Bentham's 50th birthday, Amsterdam University, Amsterdam.

Stroop, J. R. (1935) Studies of interference in serial verbal reactions. Journal of Experimental Psychology, 18, pp.187-205.

Wason, P. C. & Johnson-Laird, P. N. (1972), Psychology of Reasoning: Structure and content, Harvard University Press, Boston.

Wason, P. C. & Shapiro, D. (1971), Natural and contrived experience in a reasoning problem. Quarterly Journal of Experimental Psychology, 23: 63-71.

Wason, P. C. (1966), Reasoning, In B. M. Foss (Ed.) New horizons in psychology. Harmondsworth, UK; Penguin.