Glady, Y., **French, R. M.**, and Thibaut, J. P. (2016). Comparing competing views of analogy making using eye-tracking technology. *Proceedings of the Thirty-eighth Annual Meeting of the Cognitive Science Society*, Austin, TX: Cognitive Science Society, 1349-1354.

Comparing competing views of analogy making using eye-tracking technology

Yannick Glady, Robert French, & Jean-Pierre Thibaut
yannick.glady@gmail.com, {jean-pierre.thibaut, robert.french}@u-bourgogne.fr
LEAD-CNRS, UMR 5022, Univ. Bourgogne Franche-Comté, Pôle AAFE
21065 DIJON. FRANCE

Abstract

We used eye-tracking to study the time course of analogical reasoning in adults. We considered proportions of looking times and saccades. The main question was whether or not adults would follow the same search strategies for different types of analogical problems (Scene Analogies vs. Classical A:B:C:D scene version of A:B::C:D). We then compared these results to the predictions of various models of analogical reasoning. Results revealed a picture of common search patterns with local adaptations to the specifics of each paradigm in both looking-time duration and the number and types of saccades. These results are discussed in terms of conceptions of analogical reasoning.

Keywords: Analogical reasoning; eye tracking; analogy tasks; strategies

Introduction

Analogical reasoning is a central feature of human cognition (Gentner & Holyoak, 1997; Holyoak, 2012; Hofstadter & Sander, 2013) and involves the transfer of relations from a source domain to a target domain. Analogical reasoning has been extensively studied from adult experimental, developmental and modeling perspectives and several general models have been proposed in order to characterize this form of reasoning in both children and adults (see French 2002; Gentner & Forbus, 2011, Holyoak, 2012).

The present work uses eye tracking data to study the temporal organization of the search for a solution in adults.,. Specifically, we compare different types of analogical reasoning tasks (Scene analogies, a standard ABCD task, a Scene version of the ABCD task -- see Figures 1a, 1b, and 1c, respectively), and the search profiles resulting from participants' analysis of the stimuli for each of them. In eye tracking studies, it has been shown that there is a correlation between the amount of attention paid to a particular item and the gaze-fixation (Deubel & Schneider, 1996; He & Kowler, 1992), and between the fixation time associated with a given item and its informativeness (Nodine, Carmody, & Kundel, 1978). All of this argues in favor of using eye tracking technology to study analogy-making strategies. Indeed, any analogical reasoning task involves analyzing the stimuli, comparing them, and mapping in order to find the relations shared by the two compared domains, this according to a temporal sequence.

Existing models of analogy make different predictions regarding *how and when* participants focus on and compare stimuli. Gentner and Forbus (2011) distinguish “align-first” models from “projection-first” models. Markman and Gentner (1993) propose an “alignment-first” conception in which one first aligns the stimuli that compose the base and the target domains. From the comparisons of local elements, of local and global structures from both sides, one derives which elements should be put into correspondence (e.g., Falkenhaimer, Forbus, & Gentner, 1989). In the A:B::C:D paradigm, one would systematically tend to align A with C and look for a D (or Ds) to be aligned with B. In a scene analogy task (see Markman & Gentner, 1993; see below Figure 1), one would predict numerous saccades between the two pictures during the entire trial.

By contrast, “projection-first” models (e.g., LISA, Hummel & Holyoak, 1997) begin by searching the relation unifying the base pair (i.e., the A:B pair in the A:B::C:D paradigm) and, then, try to find matches corresponding to this relation in the target pair (i.e., the C:Ds). This model would, therefore, predict more initial attention to the A:B pair, meaning more A-B transitions and more gaze time in the first image in the Scene analogy task. Later in the trial, one expects more gazes towards C and Target and more C-Target transitions in both the A:B:C:D task or the scene analogy task.

Other strategies which have features in common with the projection-first, alignment-first distinction have been described, such as the distinction between constructive matching versus eliminative matching (Bethell-Fox, Lohman & Snow, 1984). In the constructive-matching strategy, participants concentrate on the first part of the problem before studying the second part or the solution set, which is analogous to the projection-first. In the eliminative strategy, the source and the target are compared until the best option is selected in the solution set. This strategy makes no strong prediction on the time course of a trial since participants successive elimination could be done by projections or alignments. However, there should be many transitions between A:B and C:-solutions or between scenes from the start. In any case, all these strategies rely on multiple, successive comparisons which must be coordinated.

A limited number of eye tracking studies involve analogy-making (e.g., Bethell-Fox, 1982; Gordon & Moser, 2007; Thibaut et al., 2011; Thibaut & French, 2016). Bethell-Fox et al. (1984) used an A:B::C:D task, with easy and difficult *geometrical* analogies. They found that participants with less fluid intelligence relied more on the elimination of implausible answers, i.e. had more transitions within pictures in the solution set. Interestingly, difficult items elicited more saccades back to A and B, i.e., more time spent on the A:B pair of the problem before looking to alternatives. Participants also looked at the alternatives more often than in simple trials. The authors also found that when participants first looked at the correct answer, they later tended to look at a lower number of alternatives than when an incorrect answer was first looked at. If participants looked less at other options when their first look was for the correct answer, this suggests that they had already constructed a solution for the A:B pair which allowed them to recognize that the correct solution was correct.

In developmental studies Thibaut et al. (2011) and Thibaut & French (2016) with A:B::C:D semantic analogies found key differences between adults and children. First, adults focused on the A:B pair at the beginning of the trial, paying little or no attention to C and to stimuli in the solution set. At the end of the trial, the Target was their sole focus of attention. By contrast, children organized their search around C which they actively focused on during the entire trial. At the beginning of the trial they paid more attention to C and B but also looked at the Target and the semantic distractor earlier than adults. Results also showed that children paid less attention to A and B. Overall, adults behaved in a projection-first way whereas children followed neither a projection-first (or constructive) strategy nor an alignment-first strategy but, rather, a strategy organized around C (Glady, Thibaut, French, 2012; Glady et al., 2014; Thibaut & French, 2016).

Gordon and Moser (2007) used scene analogies (see Markman, & Gentner, 1993) in which participants had to point which item in a scene had the same role as an item pointed to by the experimenter in the other scene (e.g. pointing to a boy chasing a girl, if the experimenter pointed to a dog chasing a cat, see below Figure 1). Adults initially focused on the "actor-patient" pair in the source image (i.e., a dog chasing a cat, which is analogous to our A and B terms) and then looked for the solution in the target image (a second actor-patient pair, e.g., a girl chasing a boy, analogous to our C and D terms). This is consistent with the constructive view or the projection-first conceptions (study A:B then analyze C:D). They also showed that there were longer fixations on the stimuli with an arrow pointing to it in the source scene and its match in the target scene.

Goals and Rationale

In the present paper, we compared adults' visual strategies in three different analogical tasks, the scene analogy task used by Gordon and Moser (2007) (see Figure 1a), the standard A:B::C:D task (hereafter "ABCD task", see Figure 1b) and a modified "scene" version of the A:B::C:D task in which the stimuli were inserted in scenes (hereafter, the Scene ABCD task, Figure 1c). The resulting stimuli were similar to the Scene analogy task. In the Scene Analogy task, participants have to find what plays the same role in the target scene (e.g. the boy who is chasing a girl, see **Error! Reference source not found.**) as the element that is pointed to in the source scene (e.g., a cat chasing a mouse). In this task, the stimuli are semantically connected in meaningful scenes. By contrast, in the A:B::C:? task, stimuli are presented separately and the resulting display has no global meaning. Given the instructions, participants have to compare separated pictures in order to find the common relation. The idea behind the Scene ABCD task was to insert the stimuli in simple but meaningful scene (e.g. a bird flying close to a nest). In this case the instructions remain the same as in the standard ABCD task (find the one that goes with a designated stimulus [e.g., a dog], in the same way as the bird goes with the nest). Recent evidence suggests that grouping objects in meaningful scenes might positively affect attention toward these objects and their relations (see Humphreys et al., 2010 for a review).

The main difference between the two "scene" tasks is that in the Scene analogy task, the focus is on the identification of the role of the designated object in the relation between A and B. Once this is done, one must identify stimuli connected by the same relation in the second picture and the stimulus which plays the analogous role as the one designated in the first pair. In the Scene-ABCD, when one is gazing at the AB scene, he/she has to find the relation connecting them and then orient the search in the second image around the designated C ("what goes with C") and find the object that fulfills the same role in the other objects. One could say that there is more emphasis on "role" in the Scene analogy task than in Scene-ABCD. In the standard A:B::C:D case, role might be more straightforward since location cues make them more salient.

Questions were whether the search patterns would be equivalent in the three tasks, whether the scene ABCD task would be more similar to the scene task or the ABCD task. In terms of the above models, the question was whether a unified profile of answers would appear, such as a projection first pattern (A:B then C:D) in all cases or not, or whether we would observe more alignment first (AC and BT saccades) in the scene analogy case than in the two other tasks, because participants had to identify equivalent roles in both scenes. Recall that Thibaut et al. (2011) and Thibaut & French, 2016) found a majority of A-B then C-T saccades. Finally, since scenes are semantically integrated, they might elicit more intra-domain saccades than in ABCD task.

In what follows we will attempt to test these hypotheses experimentally. Crucially, we divided the time course of each trial into three equal time slices, which allowed us to study the evolving dynamics of problem resolution. Specifically, we used a Task (Scene Analogy, Standard A:B::C:?, Scene ABCD) x Time slice (1^{st}, 2^{nd} and 3^{rd} slices) x Stimulus Type (A, B, etc.) design as within-subject factors. Response accuracy, reaction times and eye movements were recorded.

Methods

Participants

Subjects were 20 adults (14 females, 6 males; mean age=20.4 years; SD=2.21; range: from 17 to 27 years), Univ. Bourgogne Franche-Comté students. They received course credit for their participation.

Materials

Three tasks were used, each composed of three training trials and four experimental trials. The first task was a Scene Analogy task (Figure 1a), the second a Standard ABCD task (Figure 1b) and the third a Scene-ABCD task, i.e., a standard A:B::C:? task with the items composing the problems inserted in scenes (Figure 1c). In each task, the stimuli were composed of 7 black and white line drawings. The tasks were presented sequentially and appropriately counterbalanced. The order of the trials within each task was random.

In the Scene Analogy task (Fig. 1a), the pictures were based on materials in which the distractor was chosen to be semantically related to one member of the relation in the bottom picture (see Richland et al., 2006; Gordon & Moser, 2007), for example, a ball in the "boy chasing a girl" case.

In both the Scene Analogy task (Figure 1a) and the Scene A:B::C:D task, trials were constituted of two scenes (501x376 pixels each) containing 7 black and white line drawings (corresponding to A, B, C, T, semantic distractor and two unrelated distractors) framed by a black rectangle. The top scene contained the A and B pictures; the remaining pictures were in the bottom scene. The only difference in the general presentation of the two tasks was that an arrow pointed to the B stimulus in the Scene Analogy task, and to the C stimulus in the Scene ABCD task.

In the Standard ABCD trials, the A, B, C drawings were presented above the picture. On the top right, a black empty square was the solution location. The four remaining pictures (Target, Semantic Distractor and two Unrelated distractors) were presented at the bottom of the screen (Figure 1). In the analyses presented below, the potential-answer items were designated as follows: T (correct Target stimulus), SemDis (semantic distractor), and UnDis (distractor semantically unrelated to either A, B or C). The size of each picture was 200x195 pixels. Pictures were not aligned in two rows in order to minimize saccades of participant's eyes involving stimuli "on the way" to another stimulus.

The tasks were displayed on a Tobii T120 eye-tracker with a 1024x768 screen resolution using an E-Prime© software (version 2.8.0.22) embedded in a Tobii Studio (version 2.1.12) procedure to record participants' scanpaths.

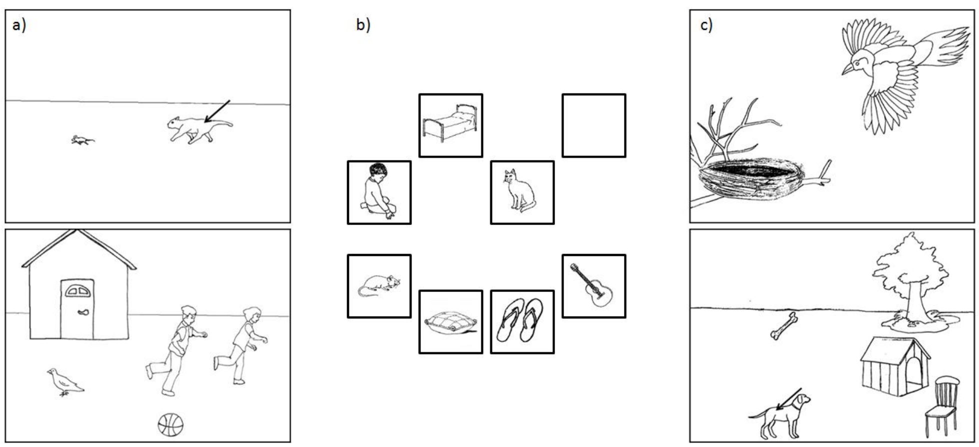

Figure 1: The three tasks used in this experiment: a) scene analogy task, b) standard ABCD task, c) Scene ABCD task.

Procedure

Participants were tested individually in a quiet room at the University. Following calibration, participants were shown a training trial and were given the instructions corresponding to each condition: In the Scene Analogy task, they were shown a stimulus in the above frame and were asked to point to the stimulus that played the same role in the below picture as the one played by the object that was pointed to in the first picture. In the Scene-ABCD task they were told "There are two pictures [pointing to A and B]. A:B go together well. Can you see why [A and B] go together?" Once the participant had given a relation linking A and B, the experimenter gave feedback. "OK! Now, do you see this one [pointing to C]? What you have to do is to find in this picture [pointing to the below scene] the item that goes with this one [C] in the same way as this one [B] goes with this one [A]." In the standard ABCD task, participants were asked to find in the solution set the stimulus that goes with C in the same way as A goes with B. No further instruction and feedback were given during test trials. Eye tracking data were recorded when the presentation of the problem started and stopped when an answer was given.

Results

Trials were excluded of the eye-tracking analysis when more than 50% of the eye movement data missing. No participant was excluded of the statistical analyses.

Participants' performance was above 95% correct in the three tasks. These analogies were simple because they were used for comparison with children in another experiment.

An ANOVA on reaction times as a function of Task Type revealed no main effect of task, $p > .2$)

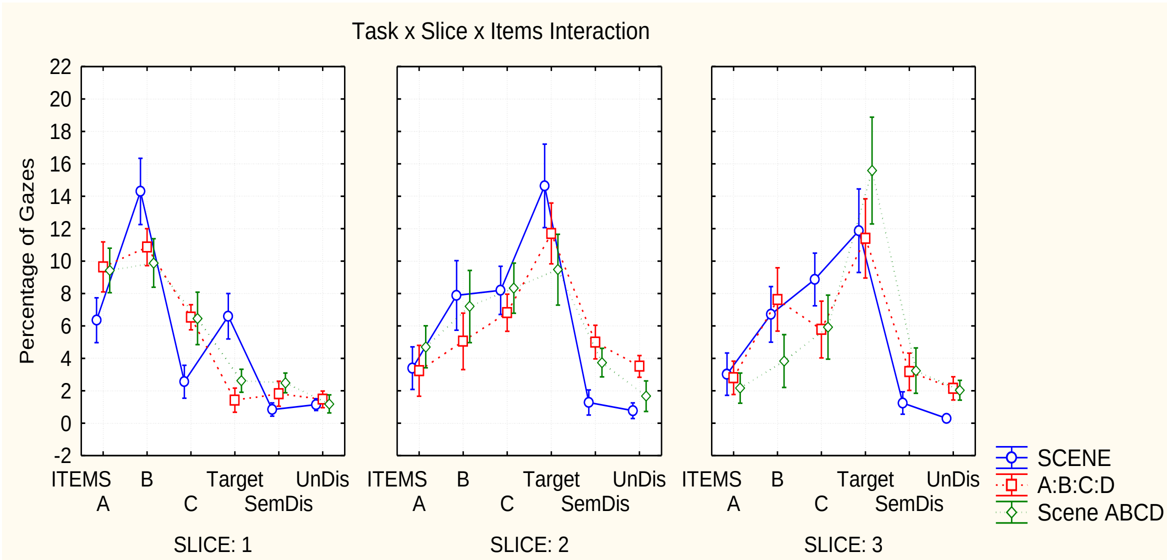

Figure 2: Interaction between Task(Scene, Standard ABCD, and Scene ABCD), Time Slice, and Stimulus Type for the percentage of gazes. Patterns of gazes were partially different in Scene condition and the two versions of the ABCD task (see text).

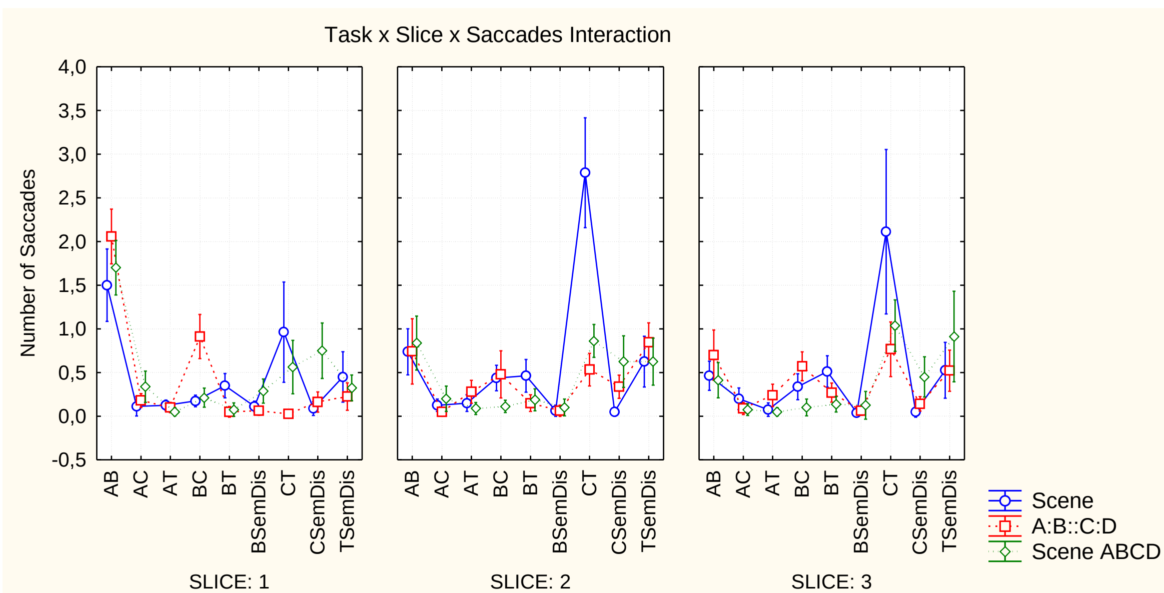

Figure 3: Interaction between Task (Scene, Standard ABCD, and Scene ABCD), Time Slice, and Stimulus Type for the percentage of gazes. Patterns of switches were partially different in Scene condition and the two versions of the ABCD task (see text).

Eyetracking data

Visual Strategies

We analyzed the proportions of gazes towards each type of stimuli with a three-way ANOVA with Type of Stimulus (A, B, C, T, SemDis, UnDis), Task Type (Scene Analogy, ABCD, Scene ABCD) and Slice (1^{st}, 2^{nd}, 3^{rd}) as within-subject factors (**Error! Reference source not found.**). There was a main effect of Type of stimulus, $F(8,328)=17.1$; $p<.001$; $\eta^2_p=.294$). There was a significant interaction between Task and Type of Stimulus, $F(10,190)=9.96$; $p<.0001$; $\eta^2_p=.34$) between Slice and Type of Stimulus $F(10,190)=3.57$; $p<.0001$; $\eta^2_p=.65$), and a significant interaction between the three factors, $F(20,380)= 6.34$; $p<.0001$; $\eta^2_p=.25$) which was the main result to be considered here.

Figure 2 shows that the three tasks started with gazes on the AB pair, which is in favor of the projection-first strategy. The Standard ABCD and the Scene ABCD tasks were quite similar in the first slice of the trial, with A and B dominant, and far fewer gazes to the potential solutions (Target, semantic distractor, and unrelated distractor). By contrast, in the Scene condition, B (i.e., the stimulus pointed to by the experimenter) was dominant at the beginning of the trial, whereas A and Target received approximately the same number of gazes, which meant that the Target received more gazes than in the two other conditions. In the second slice, C and the Target dominated in all three tasks, showing that participants were making their decision while comparing these two stimuli (see the "Saccades" section below). During this second slice, they continued to look at B (recall that in the scene analogy task, B corresponds to the object that is designated by the arrow). The third slice is similar. In the two ABCD tasks, the second slices were progressively organized around the Target. The Semantic distractor and the unrelated distractor received more gazes than in the Scene condition. This suggests that, in the ABCD tasks, participants look at the entire space of stimuli while making their decision. Note that the pattern of gazes for the ABCD task was similar to the one obtained by Thibaut and French (2016).

Saccades We also analyzed the saccades between stimuli. They tell us which stimuli are compared and when. We focused on a subset of 9 transitions, the most relevant here (see Thibaut et al. 2011; Thibaut & French, 2016). The subset was composed of A-B, C-T(arget), A-C, B-T and also A-T, B-C, B-SemDis, C-SemDis, and T-SemDis.

The first four of the above transitions are crucial to determining whether participants follow projection-first, constructive strategies (A-B then C-Target), or alignment-first strategies (A-C and B-T) or a combination of both, depending on the moment of the trial. The other transitions refer to other comparisons between C and the solution set (or within the bottom scene). We ran a three-way repeated-measures ANOVA with Transitions (A-B, C-T, A-C, B-T, A-T, B-C, C-SemDis, T-SemDis), Slice (1st, 2nd, 3rd), and Task (Scene, ABCD, and Scene ABCD) as within-subject factors. The most important result was the significant interaction between Transition Type, Task and Slice, $F(32,608)=3.40$; $p<.0001$; $\eta^2_p=.15$ (see Figure 3). The main effects and the two-way interactions were also significant.

In *the first slice*, the main common result is the presence of A-B transitions and the absence of A-C and B-T transitions. The main difference between the three tasks was on B-C which appeared only in the Standard ABCD task. One interpretation of this result is that the B and C are spatially close. Once they have received the instructions (i.e. “what goes with C”) they looked at it, and moved to A and B which they compared as witnessed by the number of A-B saccades. In the Scene analogy task, in the first slice, A-B and C-T saccades were present, suggesting that participants analyzed A and B and C-T in parallel in order to find the equivalent stimuli in the two scenes. By contrast, the search is more sequential A-B then C-T in both ABCD task with an overwhelming prevalence of A-B saccades at the beginning. The Scene-ABCD shows the same prevalence of A-B saccades, however, one clearly sees the construction of a semantic space in the second scene with the presence of C-T, C-SemDis and T-SemDis saccades. In the second slice, C-T saccades are overwhelmingly dominant in the Scene condition which suggests that participants have identified the stimuli that are involved in the target relation and are comparing them to attribute a semantic role to them. In the Scene-ABCD task, C-T C-SemDis and T-SemDis are equally distributed and there is a slight dominance of T-SemDis in the ABCD task suggesting that comparisons between semantically-related-to-C stimuli are important at this point (confirming Thibaut & French 2016). These patterns suggest that in both the ABCD and Scene-ABCD tasks participants compare C, T and SemDis in order to find the correct solution. From a theoretical point of view, evidence for A-C and B-T saccades is scarce in the three tasks, suggesting that *aligning the images is not a primary objective in any of the three tasks*, which are all organized around the A-B, C-T pattern. The pattern of results here is consistent with the projection-first hypothesis or the constructive-matching hypothesis, rather than with the alignment-first view or response-elimination strategies. Response elimination would be consistent with more saccades between the source and the target stimuli.

Discussion

The present paper extends the results of previous eye-tracking studies by comparing three different analogy tasks by means of eye tracking. Our first purpose was to assess whether and when participants tended to use projection-first or alignment-first strategies (see Introduction). Data confirmed previous studies by Thibaut and colleagues showing higher rates of A-B and C-T saccades in participants' patterns of visual search than of A-C and B-T saccades at any point during the trial, and a prevalence of gazes towards A and B in early steps of the trial. Thus the first main result is that *participants tend to construct or interpret an analogy by comparing the relation constructed in the AB domain before looking applying it in the target domain*.

A priori, it could have been argued that scene analogy trials would require more alignments (A-C and B-T transitions) (Markman & Gentner, 1993) than ABCD tasks. Indeed, in order to establish role equivalence, participants should first identify which objects might be compared in both scenes (A,B and C, T), which means that they should saccade between these stimuli in the two pairs. However, in reality, this rarely occurred (few A-C, B-T, B-C and A-T saccades). One can observe a small number of B-T transitions in the Scene task, but only in the second and third time slices. However, their number remains low compared to A-B or C-T, transitions. When they do appear, they are not accompanied by the A-C transitions that should appear simultaneously if one wants to establish the existence of parallel search of equivalent objects between the two pairs.

Regarding differences between the tasks, participants had fewer gazes towards distractors and fewer saccades involving the semantic distractor in the Scene task than in the other tasks. Overall the two scenes tasks (i.e., Scene and Scene-ABCD) differed in the way the solution was constructed. The Scene task was organized around T from the start (T gazes and C-T saccades) and fewer of the others whereas the proportion of gazes towards C, T and SemDis and transitions involving them was initially more balanced in the Scene-ABCD condition, suggesting that participants were constructing the solution from comparisons between these stimuli. Note that other factors might also play a role, such as item difficulty (not manipulated here). Indeed, Glady, French & Thibaut (2014) showed that item difficulty was associated with more gazes towards distractors, which was interpreted as an evidence for the necessity of a deeper analysis of the stimuli in order to find the solution.

In conclusion, our results revealed general patterns in the search of a solution, together with local adaptations to the specifics of the tasks.

Acknowledgements

This research has been supported by French ANR Grant 10-BLAN-1908-01, ANAFONEX to the third author, a joint ANR-ESRC grant 10-056 GETPIMA to the second author, and a FABER grant from the Conseil Regional de Bourgogne to the first and the third authors.

References

Bethell-Fox, C. E., Lohman, D. F., & Snow, R. E. (1984). Adaptive reasoning: Componential and eye movement analysis of geometric analogy performance. *Intelligence*, *8*, 205-238.

Deubel, H., & Schneider, W. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, 36, 1827–1837.

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial Intelligence*, *41*(1), 1-63.

French, R. M. (2002). The computational modeling of analogy-making. *Trends in Cognitive Science, 6(5),* 200-205.

Gentner, D., & Smith, L. (2012). Analogical reasoning. *Encyclopedia of human behavior*, Vol. 1. Elsevier Inc.

Gentner, D., & Holyoak, K. J. (1997). Reasoning and learning by analogy: Introduction. *American Psychologist, 52*(1), 32.

Gentner, D., & Forbus, K. D. (2011). Computational models of analogy. *Wiley Interdisciplinary Reviews: Cognitive Science, 2(3),* 266-276.

Glady, Y., Thibaut, J.P., and French, R. M. (2014). Adults' Eye Tracking Search Profiles and Analogy Difficulty. *Proceedings of the Thirty-fourth Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society, 535-540.

Glady, Y., Thibaut, J.P.and French, R. M. (2012). Explaining children’s failure in analogy making tasks: A problem of focus of attention? *Proceedings of the Thirty-fourth Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society, 384-389.

Gordon, P. C., & Moser, S. (2007). Insight into analogies: Evidence from eye movements. *Visual Cognition*, *15*(1), 20–35.

He, P., & Kowler, E. (1992). The role of saccades in the perception of texture patterns. *Vision Research*, *32*(11), 2151–2163.

Hofstadter, D. R. and Sander, E. (2013). *Surfaces and Essences: Analogy as the fuel and fire of thinking.* NYC, NY: Basic Books.

Holyoak, K. J. (2012). Analogy and relational reasoning. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford handbook of thinking and reasoning*. New York, NY: Oxford University Press.

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review*, *104*(3), 427.

Humphreys, G. W., Yoon, E. Y., Kumar, S., Lestou, V., Kitadono, K., Roberts, K. L., & Riddoch, M. J. (2010). The interaction of attention and action: From seeing action to acting on perception. *British Journal of Psychology*, *101*(2), 185–206

Markman, A. B., & Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive Psychology*, *25*(4), 431-467.

Nodine, C. E., Carmody, D. P., & Kundel, H. L. (1978). Searching for Nina. In J. Senders, D. F. Fisher, & R. Monty (Eds.), *Eye movements and the higher psychological functions*. Hillsdale, NJ: Erlbaum.

Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children’s development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology*, *94*(3), 249-273.

Thibaut, J. P., & French, R. M. (2016). Analogical reasoning, control and executive functions: A developmental investigation with eye-tracking. *Cognitive Development*, *38*, 10-26.

Thibaut, J.-P., French, R. M., Missault, A., Gérard, Y., & Glady, Y. (2011). In the eyes of the beholder: What eye-tracking reveals about analogy-making strategies in children and adults. *Proceedings of the Thirty-third Annual Meeting of the Cognitive Science Society* (pp. 453–458).