Towards a Visual Remote Associates Test and its Computational Solver

Ana-Maria Olteţeanu*, Bibek Gautam**, and Zoe Falomir*

*Cognitive Systems, Universität Bremen
Bremen, Germany
**Department of Computer Science,
Southeastern Louisiana University
{amoodu,zfalomir}@informatik.uni-bremen.de

Abstract. The Remote Associates Test (RAT) is a test used for measuring creativity as relying on the power of making associations, and it normally takes a linguistic form (i.e., given three words, a fourth word associated with all three is asked for). The aim of this paper is towards generalizing this test to other domains, checking for its possible application in the visual domain (i.e., given three images, an object associated to them is asked for). A pilot visual version of the Remote Associates Test (RAT-V) was created and given to human participants. A previous solver of the compound linguistic Remote Associates Test (comRAT-C) was adapted to become a prototype which can solve the visual Remote Associates Test (comRAT-V).

Keywords: Remote Associates Test, Human Creativity, Visual Associates, Computational Creativity, Cognitive Systems

1 Introduction

Humans are capable of creativity across a wide variety of tasks and domains, including the linguistic (e.g. riddles, novels), visual (e.g. visual arts, design), auditory (e.g. musical), tactile (e.g. fashion and fabrics, texture), gustatory and olfactive (e.g. culinary creativity, perfumery), etc. Creativity in many domains runs across various sensory or linguistic modalities (e.g. literature, scientific discovery, innovation).

Complex creativity tasks, like the solving of insight problems, might ellicit both linguistic and visual creativity. Creativity tests which include both visual and linguistic elements do exist - like the Torrance Tests of Creative Thinking (TTCT), which contains both verbal and figural tests [6]. However, no such tests exist which can be given separately in both linguistic and visual forms, thus affording cross-domain comparison of a particular set of creative processes. The usefulness of such a test would be to: (i) check whether the same creative processes act across domains; (ii) compare performance results in various domains; and (iii) posit domain-relevant differences.

Aiming to fill this gap, this paper takes a well established creativity test, the Remote Associates Test [7] – for which a previous computational linguistic solver was implemented (comRAT-C [10]) under a theoretical creative problem-solving framework (CreaCogs [9, 8]) – and describes our approach towards developing a visual derivate of this test.

The rest of the paper is organized as follows. The Remote Associates Test and the construction of its visual counterpart (vRAT) are discussed in Section 2. A study with human participants who were given vRAT queries is described in Section 3. A short description of the linguistic comRAT-C together with its current prototype adaptation to solve visual queries is discussed in Section 4. Results on the experimentation carried out with human participants are provided in Section 5, while results of the computational comRAT-V prototype are described in Section 6. A discussion of this pilot test and prototype system are provided in Section 7 and further work is proposed.

2 Outlining the Remote Associates Test (RAT) and its Visual Counterpart

Imagine you are given three words - like CREAM, SKATE and WATER - and asked which is a fourth element common to all of them. This describes the Remote Associates Test originally devised by Mednick and Mednick [7]. The answer to this particular query is ICE.

The Remote Associates Test has been used in the literature [1, 5], and adapted to various languages [2, 4]. To check whether this creativity test could be adapted to more than linguistic examples, the authors decided to work towards a visual version of the RAT.

Different versions of the RAT [7] exist, after some researchers have argued that the items in the test were not all equal. Worthen and Clark [11] argued that some of these items are functional, and others structural. Functional items are those between which a non-language relationship is present (e.g. items like "bird" and "egg"), while structural items have previously been associated within a syntactic structure (e.g. items like "black" and "magic"). Compound remote associates correspond to structural associates in Worthen and Clark's categorization.

Normative data from compound remote associates [3] has been used before by the authors to evaluate a computational solver of the RAT [10] implemented using language data. In this paper, the authors use their understanding of this task to build a visual Remote Associates Test. In a previous formalization [10], the Remote Associates Test was described as follows: 3 words are given, w_a, w_b, w_c, and a word needs to be found, w_x, which relates to all three initial words. In the compound RAT case, terms $(w_a, w_x), (w_b, w_x)$ and (w_b, w_x) or their reverse, $(w_x, w_a), (w_x, w_b), (w_x, w_c)$ have to be successive or composed terms in the language in which the RAT is given in. In the case of composed terms, w_z might be another word composed of one of the initial terms and the solution term, like $(w_x w_a)$ or $(w_a w_x)$. For example, for the query AID, RUBBER and WAGON, the answer term BAND constructs composed terms with some of the query terms (BAND-AID, BANDWAGON), but not with others (RUBBER BAND). Note that the answer term is also not in the same position in the three linguistic structures.

In order to devise a visual RAT, the same mechanism was applied, with entities w_a, w_b, w_c and w_x being visual representations of objects and scenes. Thus, given entities w_a, w_b, w_c, there exists an entity w_x, which generally co-occurs visually with the other shown entities w_a, w_b and w_c.

For example, Fig. 1 provides the following entities: HANDLE, GLOVE and PEN. HAND is an appropriate answer to this query, being a visual entity which co-occurs with each of the given three. The visual entity HAND can be considered a visual associate of each of the initial objects HANDLE, GLOVE and PEN.

Fig. 1. Example of a visual RAT question. This is the first training query, showing the participants the following visual entities: HANDLE, GLOVE and PEN.

Each initial object is considered to have a variety of other visual associates. Therefore, this work assumes that visual associates are terms which play the role that word terms play in the language-based RAT. Visual associates which co-occur together, in a previously encountered visual scene or experience, play the same role as composed words or linguistic structures in which w_a and w_x co-occur. Thus, visual experiences containing the visual entities (HANDLE, HAND), (HAND, GLOVE) and (PEN, HAND) are required to solve the visual query shown in Fig. 1.

Next section explains the visual RAT test carried out by human participants.

3 Study with Human Participants: Answering the Visual RAT and Providing Visual Associates

The study carried out on human participants contained two parts. Participants where asked: (1) to solve some visual RAT queries and (2) to provide visual associates to some concepts not included in the previous queries. Participants were split in 4 groups, each group being given part of the RAT queries to solve, and the objects in the other queries to provide visual associates for.

Part 1

20 visual RAT queries plus 2 initial examples were set-up for initial experimen-

tation with human participants. Each of the queries showed 3 visual stimuli: objects (5 in training, 54 in test) or scenes[1] (1 in training, 6 in test).

The given training examples are showed in Fig. 1 and Fig. 2. The answer item is not contained in either of the initial images. The expected process is that participants could ellicit their visual memory about such co-occurrences of visual associates. Note that, Fig. 2 avoids presenting the sea while presenting the image of a beach, as the expected answer to this visual RAT query is WATER.

Fig. 2. The second training vRAT query showed the items above to the participants: BATHTUB, GLASS and BEACH.

Participants were instructed that:

- they would be presented with three objects or scenes, and asked to find a fourth element that is related to each of them;
- they could then choose between various ways in which they first perceived the answer when they arrived at it: (i) Visual imagery (they imagined the answer), (ii) Word (they thought of the answer verbally) and (iii) Other (in this case, they were asked to specify);
- they should provide a difficulty rating for each test item on a Likert scale, with a range from 1 (Very Easy) to 7 (Very Hard).

Afterwards, the test with the visual RAT queries followed.

Part 2

Participants were asked to contribute visual associates to a set of objects, which were query items for queries they have not received, as explained before. This task was explained as follows:

> *Visual associates are things you see when you imagine a particular object. These might be other objects, which are situated next to the object that you are imagining in some circumstance, or specific parts of the object you are imagining.*

[1] A *scene* is considered a visual display in which multiple objects might be considered salient. Parts of other objects may also be present when showing an *object* entity, but these parts were clearly not salient stimuli.

For example, visual associates for "glove" might be: hand, thorns, snow, scalpel, hot pan, bike, dirt. Visual associates for "pen" might be: paper, notebook, letter, test, form, cheque, desk, ink, drawing, writing, pen holder, ear, pen case, pencil, etc.
Imagine each item, and then write the visual associates that come to mind.

Grouping Procedure
The test was administered to four groups, via four different surveys developed using Google forms. The participants were asked to select their group themselves using a randomizer[2] which presented two Euro coins, on head or tails position. Depending on the coins arrangement provided by the randomizer, participants proceeded to one of the four groups tasks. All groups were shown the same initial two training examples. The 20 questions were split in four 5-question groups. Each of the four groups was asked to solve 3 sets of questions (thus 15 vRAT queries), and asked to offer visual associates for the objects in the fourth group of queries (thus 15 objects). The types of tasks (questions + visual associates) given to each group are specified in Table 1.

Table 1. The four groups in the study and their assigned tasks. Note that "Q" denotes a question, and n the number of participants in each group.

Study items	**Group 1** $n = 8$	**Group 2** $n = 15$	**Group 3** $n = 8$	**Group 4** $n = 12$	**Answers per item**
vRAT Training Examples	Yes	Yes	Yes	Yes	Shown to all
vRAT Q1-5	Yes	Yes	Yes	**No**	Gr. 1, 2, 3 ($n = 31$)
vRAT Q6-10	Yes	Yes	**No**	Yes	Gr. 1, 2, 4 ($n = 35$)
vRAT Q11-15	Yes	**No**	Yes	Yes	Gr. 1, 3, 4 ($n = 28$)
vRAT Q16-20	**No**	Yes	Yes	Yes	Gr. 2, 3, 4 ($n = 35$)
Visual associates for objects in questions	Q16-20	Q11-15	Q6-10	Q1-5	all objects across groups

Note that participants did not provide visual associates to a vRAT test item that they have previously answered, in order to avoid bias towards mentioning associations which were already made salient by the test items. The design we used in this study allowed for visual associations to be given to all objects across participants.

4 A Visual Computational Solver (comRAT-V)

This section describes how the computational visual RAT problem-solver works by describing its knowledge base content (Section 4.1) and its query solving process (Section 4.2).

[2] `https://www.random.org/coins/?num=2&cur=60-eur.germany-1euro`

4.1 Knowledge Base in comRAT-V

A previous system, comRAT-C, solved the compound RAT using language data [10]. Specifically, the most frequently occurring words appearing together as a tuple (2-grams or bigrams) were obtained from a genre-balanced Corpus of Contemporary American English (COCA)[3].

As the authors could not find in the literature any visual linked and annotated database which included the concepts used in the 20 queries included in the human test, the strategy followed was to ask the participants in the study for visual associates, as the previous section explains. Therefore, visual associates were obtained for all objects appearing in the 20 vRAT queries, that is, participants provided visual associates for a total of 60 objects. The objects were presented in such a way that a common associate will not be salient. These visual associates where used for the Knowledge Base (KB) of comRAT-V in the same way in which 2-gram relations were used by comRAT-C.[4]

Data thus obtained was cognitively valid data of visual associates obtained via introspection. This data was given to comRAT-V, which used it to construct its (visual) Concepts and Links knowledge base. The queries to be shown to humans were then given to comRAT-V. For each query, the three Concepts or Objects given in the query were ellicited from the KB, then Links were used to yield their visual associates. comRAT-V then offered the item(s) it converged upon as a possible answer.

A faster automatic way of extracting object associates from visual scenes data can be envisaged (see Section 7). However, the current prototype served our purpose to check whether comRAT will work with visual domain queries, and what was its performance.

4.2 Query Solving Process

The comRAT-C organized the data in its KB in Expressions, Concepts and Links between co-occurring Concepts. The comRAT-C solved RAT queries by activating the Concepts involved in each query in its KB, using the Links to navigate to syntactical items which those Concepts co-occurred with, and offering as a possible answer those items upon which this search and activation process converged, as shown in Fig.3.

The comRAT algorithm has been generalized to solve the linguistic and the visual RAT, which are equivalent in the nature of the processes they ellicit, although the type of data they input is different. Thus the likelihood of finding an answer based on frequency of the known items is computed in comRAT-V as in comRAT-C [10] when the system needs to choose one of multiple possible answers. When no 3-item convergence is made, comRAT-V checks for 2-item

[3] Corpus of Contemporary American English (COCA): `http://corpus.byu.edu/coca/`

[4] Thus an object and its visual (and implicitly spatial) associate is considered to be similar to a language term and its syntactic neighbour.

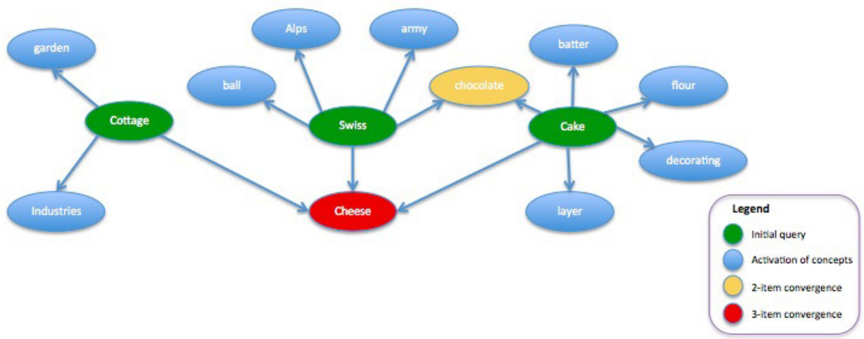

Fig. 3. Visual depiction of the search and convergence process in comRAT-C.

convergences. If multiple such terms are found, comRAT-V proceeds to compute the most likely of the terms and offers it as an answer.

5 Results from the Visual RAT Test with Human Participants

This section describes the participants to the visual RAT and vRAT results.

Participants

43 participants completed the study, 30 male and 13 female. The ages of the participants ranged between (btw.) 20 and 60 years old (y.o.), as follows: 6 btw. 20-30 y.o, 19 btw. 30-40 y.o., 14 btw. 40-50 y.o., 4 btw. 50-60 y.o. The self-assessed English level of the participants ranged between Intermediate and Native, as follows: 9 Intermediate, 21 Advanced, 10 Proficient, 3 Native.

Results

As shown in Figure 4, the percentage of participants solving the set of queries varied, between 6.45% (Q5) and 97.1% (Q20), with an average query solving percentage of 63%. Based on this, some queries may be classiffied as the three most difficult (Q5, Q13, Q16) and others as the three easiest (Q8, Q18, Q20).

As shown in Figure 5, participants declared they first perceived the answer mostly visually (56.6%) or as a word (38.9%). Some participants also declared that they did not know (3.26%) or that they perceived the answer via another sense, like feeling the heat when the answer was fire (0.16 %).

6 Results of the Computational Visual RAT (comRAT-V)

Visual associates provided by the participants to our study were added to comRAT-V's knowledge base. With this data, and no use of query frequency comRAT-V was already solving 14 of the 22 query items (63.64%). Then comRAT-V calculated the frequency of occurence of the visual associates, in order to apply

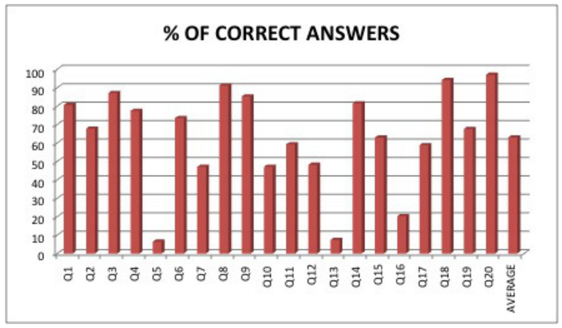

Fig. 4. Percentage of correct answers per query, as solved by the human participants.

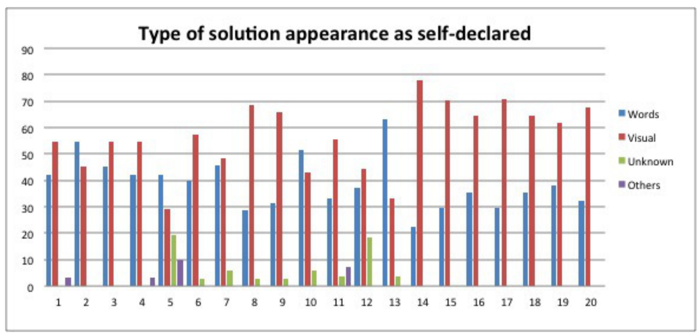

Fig. 5. Type of solution appearance as self-declared by participants.

the same frequency-based likelihood algorithm as comRAT-C [10] when selecting the answers. Given this data, comRAT-V managed to answer correctly 16 out of the 22 items, as shown in Table 2. Out of these, 13 correct answers came from 3 known items convergences, with 3 answers coming from 2 items known convergences.

Table 2. Analysis of the accuracy of responses provided by the system.

	1 item known	2 items known	3 items known	Total
Correct	0	3	13	16
Plausible	0	3	0	3
Not solved	2	1	0	3
Total	2	7	13	22
Accuracy	-	42.86% (85.71%)	100%	72.73%

Some queries encountered two or more possible answers. For Q8, two answers are possible from a 3-known items convergence - MEAT and CHEESE. However, the correct answer, MEAT, is chosen due to the frequency based likelihood. Similarly, Q21, in which a COMB, RAZOR and SHAMPOO are presented, encounters a larger set of possible correct answers. Amongst the possible three-item convergence answers (e.g.WATER, BATHROOM, MIRROR, etc.), the correct answer HAIR was chosen by comRAT-V. Queries can be answered correctly based on a two item convergence - for example Q14 was answered in this way, as only two of the visual associates linked the query items to the answer.

7 Discussion and Further work

Our current visual RAT prototype showed promise, as human participants were able to solve it (63%), a variety of difficulties were present in the different queries and 56.6% of participants said they arrived at the answer through visual imagery. Moreover, various participants declared that they enjoyed the vRAT test.

Whether queries were or were not solved through a visual imagery process is yet to be proven, as subjective reports are not reliable in this case. A fMRI-based experiment showing different language-based compound queries and vRAT queries might be able to show whether this is indeed the case. Humans might still translate visual stimuli in language stimuli, especially as the answer was asked for in language, and semantic relations are hard to avoid altogether.

However, as comRAT-V performed well based on visual associates provided by the human participants, we can assume that the queries can be solved using visual associations by humans as well. More visual affordance data is required to strengthen the current results, as these are based on visual associates and frequency of visual associates provided by the participants. As further work, the authors will focus on gathering more data for the comRAT-V knowledge base. Two ways to gather such data are envisioned:

- Get more human participants to provide visual affordances to all the objects used in the vRAT test, without giving them the test and/or
- Find a way to extract such visual associates automatically from images depicting indoor and outdoor scenes.

The authors plan to analyze whether there is a relationship between the results obtained with comRAT-V and human results in the vRAT. The authors also plan to increase the number of queries for the vRAT, since a larger set of queries might provide more insight and stronger results. A future focus will also be to investigate the different classes of difficulty in such queries, the preferred answers in multiple queries and the relation between fluency in providing visual associates by human participants and ability to solve the vRAT.

Acknowledgments

Dr. Ana-Maria Olteţeanu and Bibek Gautam are grateful for DAAD's support through its RISE project. Dr.-Ing. Zoe Falomir acknowledges funding by the project COGNITIVE-AMI (GA 328763) by the European Commission through FP7 Marie Curie IEF actions. All three authors acknowledge the support offered by the Universität Bremen and the Spatial Cognition Centre.

References

1. Ansburg, P.I.: Individual differences in problem solving via insight. Current Psychology 19(2), 143–146 (2000)
2. Baba, Y.: An analysis of creativity by means of the Remote Associates Test for Adult Revised in Japanese (JARAT FORM A). Japanese Journal of Psychology (1982)
3. Bowden, E.M., Jung-Beeman, M.: Normative data for 144 compound remote associate problems. Behavior Research Methods, Instruments, & Computers 35(4), 634–639 (2003)
4. Chermahini, S.A., Hickendorff, M., Hommel, B.: Development and validity of a Dutch version of the Remote Associates Task: An item-response theory approach. Thinking Skills and Creativity 7(3), 177–186 (2012)
5. Dorfman, J., Shames, V.A., Kihlstrom, J.F.: Intuition, incubation, and insight: Implicit cognition in problem solving. Implicit cognition pp. 257–296 (1996)
6. Kim, K.H.: Can we trust creativity tests? A review of the Torrance Tests of Creative Thinking (TTCT). Creativity research journal 18(1), 3–14 (2006)
7. Mednick, S.A., Mednick, M.: Remote associates test: Examiner's manual. Houghton Mifflin (1971)
8. Olteţeanu, A.M.: Two general classes in creative problem-solving? An account based on the cognitive processes involved in the problem structure - representation structure relationship. In: Besold, T., Kühnberger, K.U., Schorlemmer, M., Smaill, A. (eds.) Proceedings of the International Conference on Computational Creativity. Publications of the Institute of Cognitive Science, vol. 01-2014. Osnabrück (2014)
9. Olteţeanu, A.M.: From simple machines to Eureka in four not-so-easy steps. Towards creative visuospatial intelligence. In: Müller, V. (ed.) Philosophy and Theory of Artificial Intelligence. Synthese Library, Berlin:Springer (to appear)
10. Olteţeanu, A.M., Falomir, Z.: comRAT-C - A computational compound Remote Associates Test solver based on language data and its comparison to human performance. Pattern Recognition Letters (2015), http://dx.doi.org/10.1016/j.patrec.2015.05.015
11. Worthen, B.R., Clark, P.M.: Toward an improved measure of remote associational ability. Journal of Educational Measurement 8(2), 113–123 (1971)