

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Structural vs. Superficial Similarity During Unprompted Analogical Retrieval: Which one Exerts a Greater Force?

Permalink

<https://escholarship.org/uc/item/10n7010j>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors

Minervino, Ricardo A
Trench, Maximo

Publication Date

2022

Peer reviewed

Structural vs. Superficial Similarity During Unprompted Analogical Retrieval: Which one Exerts a Greater Force?

Ricardo A. Minervino (MINERVINORA@Gmail.Com)

Department of Psychology, Universidad Nacional del Comahue
IPEHCS-CONICET/UNCo

Máximo Trench (MAXIMO.TRENCH@Crub.Uncoma.Edu.Ar)

Department of Psychology, Universidad Nacional del Comahue
IPEHCS-CONICET/UNCo

Abstract

Traditional laboratory studies have found that people are more likely to retrieve surface matches than distant analogs, suggesting that superficial similarities exert a stronger influence than structural similarities on retrieval. However, it has been contended that the observed supremacy of surface similarity may have originated in experimental conditions that are unfairly adverse for the retrieval of distant analogs, as well as in a faulty separation between surface and structural similarity during the construction of surface matches. In two experiments, we presented a target item that maintained only superficial similarities with one extra-experimental source and only structural similarities with another one. By using natural items, we were able to avoid the shallow processing often attributed to experimental analogs, while carefully controlling that surface matches did not maintain structural similarities. Converging with traditional results, our data showed a more frequent retrieval of surface matches than of distant analogs, indicating a supremacy of superficial similarities during retrieval.

Keywords: analogy; retrieval, superficial similarity

Introduction

The ability to draw analogies across thematic domains appears in all short-lists of inherently human capacities, as it subserves activities as sophisticated as scientific discovery (Hesse, 1966), argumentation (Blanchette & Dunbar, 2001), instruction (Richland, Holyoak, & Stigler, 2004) or metaphor (Lakoff & Johnson, 1980). By linking the elements of a poorly understood situation (*target analog*) to those playing parallel roles in a better-known situation (*base analog*), unmapped base elements can be projected onto the target analog. Suppose that Jenny applies for a role in a play, but after being rejected she tells her friend that it is better this way, since the script was not good enough. This reaction could perhaps bring to her friend's memory a famous fable from Aesop, wherein a fox fails to reach some grapes from a tree, and consoles itself by affirming that the grapes were sour after all. The above situations could be considered analogous to the extent that they maintain *structural similarities*, that is, they instantiate the same causal pattern of events (i.e., someone fails to achieve a goal, and then minimizes its importance). In contrast, the semantic similarities between entities playing like roles in the compared situations (e.g., actress ↔ fox, a play ↔ grapes) do not represent a requisite for an analogy to

hold, and are therefore regarded as *superficial similarities*. Just as base analogs stored in long-term memory (LTM) can sometimes maintain only structural similarities with the target, other stored items might maintain only superficial similarities. Consider Köhler's experiment wherein a chimpanzee manages to reach some bananas by piling up two boxes and stepping on them. Despite some resemblances with the sour grapes fable at the level of isolated entities and relations, both situations instantiate clearly different principles. Situations maintaining similarities only at the level of isolated relations and entities are called *surface matches*.

A question of central interest to the discussion about the efficacy of our cognitive system concerns the kind of similarity that governs the retrieval of related content from memory. Put more simply, are we more likely to retrieve surface or structural matches to a working memory probe? The procedure typically employed to address the question of which kind of similarity governs retrieval consisted in having participants read a series of stories during a first phase, and then presenting participants with a new set of stories, each one maintaining some kind of similarity with one of the stories of the previous phase. For each of the stories of the second set, participants have to report which stories from the first set such cue reminded them of. As an example, one of the target stories used by Gentner, Rattermann and Forbus (1993) stated that an old hawk was attacked by a hunter who needed feathers for his arrows. After the hunter missed, the hawk glided down and offered to give him a few feathers, for which the hunter pledged never to shoot the hawk again. The structurally-similar base situation of this set of materials shared first-order and higher-order relations with the target, but did not maintain similarity at the level of entities: It told that a warlike country wanted its neighbor's powerful computers and attacked it with missiles. After the aggressor failed, the attacked country offered to sell some of its computers to it, with the result that its neighbor country promised never to attack it again. In contrast with the structurally-similar base item, the superficially-similar item maintained similar objects and first-order relations with the target, but not a common causal pattern of events: It told that an eagle received a failed attack from a sportsman, who tried to kill her with a crossbow and some featherless arrows. As the eagle suspected that the sportsman wanted her feathers,

she glided down and offered him a few, with the result that he promised never to attack her again. But he did, and as the eagle was falling from the tree, she realized that the arrow was steered by her own feathers. Across a series of experiments following this procedure, Gentner et al. (1993) obtained that superficial matches were retrieved in nearly 50% of the cases, whereas structural matches were retrieved in less than 20% of the cases. Later studies using this procedure (e.g., Catrambone, 2002; Warton, Holyoak, & Lange, 1996) yielded similar results, thus informing the development of computer simulations that mimic this superficial bias (e.g., Forbus, Gentner, & Law, 1995; Hummel & Holyoak, 1997).

With the turn of the century, several authors (e.g., Blanchette & Dunbar, 2000; Hofstadter & Sander, 2013; O’Keefe & Costello, 2008) have contended that the supremacy of surface similarity obtained in traditional experimental studies may have originated in experimental conditions that are unfairly adverse for the retrieval of distant analogs, as well as in a faulty separation between surface and structural similarity during the construction of experimental materials. As opposed to series of as many as 32 inconsequential stories—which participants “swallow” one after the other without much opportunity for elaboration—the base situations to which we resort when generating analogies tend to be, in general, extremely familiar. According to Hofstadter and Sander (2013), we all rely on base situations to which we have devoted our full attention, and which have been consolidated over time. In turn, their repeated retrieval is likely to elicit their generalization, increasing the likelihood that such knowledge becomes activated during the processing of superficially dissimilar analogs (Ross & Kennedy, 1990). Hence, both the experimental stories and the conditions under which they were learned might not do justice to the way in which the extra-experimental episodes that we naturally employ as base analogs are typically encoded.

On the other hand, it has been demonstrated that participants regard some of Gentner et al.’s (1993) surface items as sharing not only objects and first-order relations with the target, but some degree of higher-order structure as well (Raynal, Clement, & Sander, 2020, Experiment 1B). To illustrate, despite an obvious mismatch in how *Karla the Hawk* and *Zerdia the Eagle* end, there is a non-trivial structural commonality in that in both cases someone attempts to avoid an attack by offering the attacker the resource that the attacker is presumed to desire. As the segment that precedes its mismatching ending could serve as a structural cue to the base story, the assertion that its retrieval was based exclusively on surface similarity is problematic.

In a series of experiments using a story-reminding paradigm, Raynal et al. (2020) reassessed the superiority of surface vs. structural similarity during retrieval by means of a set of stories in which the surface matches were specially crafted to avoid any degree of structural overlap with the target. Somewhat reminiscent of *Karla the Hawk*, this target involved an ambulant pizzaiolo who held a pizza truck in a popular place, but who reacted to the settlement of another

pizzaiolo in a nearby location by giving him generous advice on how to improve his dough. The target story ended that in order to show how much he found his intention was nice, the second pizzaiolo relocated his truck to avoid competing with him. While the analogous match reinstated the abstract features of the target in the context of two girls who competed for attention from a boy to whom they were both attracted, the superficial match only shared isolated objects and relations with the target: It involved a food truck called “At Alessandro & Fabio’s”, whose clientele were fond of the authentic atmosphere steaming from this stand, held by two happy looking pizzaioli dressed in traditional Italian suits. However, the story ended that once the two pizzaioli had left this selling space, they switched to traditional German clothes for selling sausage specialties at “Hans & Hendrich’s”. Across three experiments, Raynal et al. (2020) found that the retrieval of structural matches (around 80%) was nearly four times higher than the retrieval of superficial matches, a pattern that stands in sharp contrast with that of Gentner et al. (1993). As the retrieval of surface matches could not be attributed to the concurrent presence of structural similarity, the authors argued that surface similarity does not represent the main contributor to retrieval. On this account, our cognitive architecture is reasonably well-suited for locating abstractly related items in LTM, even when they compete with purely superficial matches.

Despite the apparent superiority of Raynal et al.’s (2020) materials for assessing the supremacy of surface vs. structural similarity during the retrieval of related content from memory, resorting to a narrower set of stories (between 6 and 8) to be read during the learning phase might conceivably allow for serially comparing the target against each of the stories from Phase 1, therefore circumventing the massive computation that would otherwise be needed to search the whole of LTM. As we see it, the traditional version of the story-reminding paradigm faces a conundrum: While the inclusion of too many stories (as in Gentner et al., 1993) hinders the possibility of achieving an adequate encoding of base stories in LTM, the decision to include a narrower set of stories allows participants to execute a serial type of search that trivializes the problem of navigating LTM at large. In light of this trade-off, we reasoned that the possibility of resolving the current inconsistency of results about the kind of similarity that governs memory retrieval would benefit from moving away from the traditional 2-phase version of the story-reminding paradigm, and towards more ecologically-valid conditions wherein participants are not prompted to focus their search on an episodically-predefined set of memory items. While it is true that memory search can on certain occasions be circumscribed to an episodically-bound set of alternatives (e.g., as when needing to relate a physics problem to the problems reviewed in the chapters covered by the quiz), the most prevalent opportunities for analogical retrieval are at the same time the most challenging, in the sense that we not only lack a subset of LTM wherein to circumscribe our search, but typically lack any hint about the availability of analogous cases in the memory set. To

complicate matters more, the temporal and contextual separation between the base and the target encodings are dramatically larger than in traditional experimental studies¹.

To assess the kind of similarity that governs the spontaneous type of retrieval that takes place in these more stringent but realistic set of conditions, we adapted the hybrid procedure previously employed to compare the retrieval of naturally-encoded structural analogs presenting different degrees of surface similarity with the target while controlling for the potentially uneven availability of these types of items in LTM (Trench & Minervino, 2015). As the base situations included in the present study were selected to be especially compelling to participants, its deployment as base analogs could potentially overcome the artificiality and lack of personal engagement often ascribed to experimental stimuli used in the story-reminding tradition. For Experiment 1, we selected two popular movies seen by most of the population under study, and crafted a target situation that shared structural features with one of the movies, and superficial features with the other. Once participants had read the target situation, they were asked to report any movie that had spontaneously come to mind, even if briefly, while reading the story. The second experiment replicated this procedure with another type of personally significant extra-experimental episodes: widely reproduced public events.

Experiment 1

Method

Participants An initial sample of 422 Argentine residents were invited to volunteer for the study through email, Facebook and WhatsApp. The only requisites for participation were being over 18 years, having completed high school education, and not having taken university-level courses of Psychology. The final sample comprised 138 participants between 18 and 64 years of age ($M = 29.44$, $SD = 9.04$, 54% female) whose scores on the questionnaires presented at the end of the study showed evidence of faithful memory of at least one of the two movies employed as base items.

Materials In order to assess the spontaneous retrieval of structural and/or superficial matches encoded in extra-experimental settings, participants read the following story:

Based on images captured by a space probe, geologists from Harvard succeeded in artificially reproducing the quakes that take place in the Tethys Moon of Saturn. Their objective is to observe them more directly. These recreations are being carried out in the desert of Monument Valley, Arizona. The site has been sealed with an isolating material that precludes the propagation of the seismic activity outside of the experimentation zone. Before inviting external geologists to observe the earthquakes, the University needs approval from Arizona's Secretary of Safety Against Natural Disasters. The engineer in charge of this organization has doubts regarding the project, and has warned the geologists that the artificial reproduction of natural phenomena may engender unforeseeable results, which often escape human control. In order to assess the safety of the anti-seismic belt, the

engineer visits the site in a helicopter driven by a rescue firefighter. The belt begins to fail, and the vibrations radiate some 5km outside the perimeter. The Secretary of Safety proceeds to the cancellation of the project.

The fictional situation presented above maintains structural correspondences with "Jurassic Park", a tremendously successful adaptation of Michael Crichton's best-selling novel that had been watched by most participants of the study, sometimes more than once. The plot involves a wealthy businessman who had succeeded in reproducing dinosaurs out of fossil DNA, and who plans to use them as a touristic attraction. During an inspection of the park, scientists from different disciplines raise doubts about the safety of the project, under the argument that such complex biological phenomena might unfold in unpredictable ways. During the inspection, the dinosaurs break down the security fences and attack people, thus precipitating the cancellation of the project. The structural similarity between "Jurassic Park" and the target analog herein presented has to do with the fact that in both cases (a) there is a project to artificially reproduce natural phenomena, (b) there is an intention to exhibit such phenomena to a broader audience, (c) qualified experts raise doubts about the safety of the project, (d) during an inspection to the site, the reproduced phenomena get out of control, and (e) this outcome precipitates the cancellation of the project.

At the same time, the target situation presented to participants maintained superficial similarities with another popular movie: "St. Andreas". In this movie, a rescue firefighter drives a helicopter from Los Angeles to San Francisco accompanied by his ex-wife, with the initial objective of rescuing their daughter from an earthquake of apocalyptic dimensions. While flying over the affected areas, they help other victims that need to be rescued. This partnership aids in rebuilding their relationship, as well as in healing the loss of an older daughter of theirs some years before. Even though the plot of "St. Andreas" shares a few isolated entities and relations with our target situation (an earthquake, a geologist and a helicopter steered by a rescue firefighter), their causal patterns of events are completely unrelated.

Procedure The initial cohort of candidates were invited through email and social media to participate anonymously in a brief study on text comprehension. Those responding positively were subsequently sent a link to a Google Form, whose first page consisted in an informed consent. Upon accepting to participate, the "next" button led to a second page dedicated to collect demographic information such as age, gender, and achieved level of studies. After the demographics section, a text segment asked participants to read the subsequent material very carefully, so as to be able to answer comprehension questions about it. Participants could read the text for as long as they wanted, and had to press "next" to proceed to the following section. In the following page they were asked to report whether any movies had come to mind, even if just briefly, while reading the previous text. For those responding "yes", a space opened up to list a maximum of four movies. In

case any of the recalled movies belonged to a series (e.g., Terminator 1, 2 or 3), participants were asked to specify of which particular movie they had been spontaneously reminded. The next question asked participants whether they had seen “Jurassic Park”. In order to avoid counting as retrieval failures those cases where the critical information to be retrieved was not available in LTM, participants responding “yes” were further asked to respond 5 multiple-choice questions about the movie’s plot (see Table 1). In a similar manner, participants were queried about whether they had seen “St. Andreas”. Only participants who got at least 4 of the 5 questions in at least one of the two questionnaires were retained for further analysis.

Table 1: Questions used to assess participants' representation of the plot structure of Jurassic Park

-
1. A philanthropist succeeded in cloning dinosaurs out of
 - (a) DNA in fossilized mosquitos that had bitten dinosaurs
 - (b) fossilized dinosaur eggs
 - (c) frozen dinosaurs found in a glacier
 - (d) bones retrieved from underground layers
 2. The philanthropist aimed at
 - (a) opening the park to the visit of scientists
 - (b) cloning other extinct species beyond dinosaurs
 - (c) opening the park to the general public
 - (d) selling the dinosaurs to zoos from all over the world
 3. Scientists plus a representative of the investors
 - (a) assess the feasibility of building similar parks in other places
 - (b) visit the park to assess its safety
 - (c) travel to the park to attend its inauguration ceremony
 - (d) intend to conceal unethical procedures related to how the dinosaurs are being cloned
 4. One of the scientists expresses that
 - (a) the dinosaurs are not being adequately fed
 - (b) certain dinosaur species could attack smaller species
 - (c) technology should not be used to clone human beings
 - (d) the artificial reproduction of natural phenomena often escapes human control
 5. During the visit to the park
 - (a) dinosaurs escape their cages and the project is cancelled
 - (b) dinosaurs kill two of the three scientists
 - (c) an unknown illness kills most of the dinosaurs
 - (d) it is decided that the project could be approved pending some safety improvements
-

The “Jurassic Park” group ($n = 56$) comprised participants who passed the Jurassic Park quiz, but either had not seen St. Andreas or did not pass its quiz. In turn, the “St. Andreas” group ($n = 33$) comprised participants who passed the St. Andreas quiz, but either had not seen “Jurassic Park” or had not passed its corresponding quiz. Participants passing both quizzes ($n = 50$) were assigned to the “Jurassic Park + St. Andreas” group.

Results and Discussion

Among participants who had been exposed to either the superficial match (St. Andreas) or the structural match

(Jurassic Park) but not to both, the superficial match was more frequently retrieved than the structural match 48.7% vs. 16.1%, $\chi^2(1, 89) = 6.07, p = .0137, \phi = .26$. The central result of the present research, however, concerns the responses given by participants who had both a surface match and a structural match available in LTM. Like in the previous contrast, an analysis of the responses given by the Jurassic Park + St. Andreas group revealed that the superficial match was more frequently recalled than the structural match, 52% vs. 14%, McNemar’s $\chi^2(1, 50) = 13.37, p < .001, \phi = .52$.

The results of the between and the within-participants comparisons are consistent with those of the first generation of studies using a cued-recall paradigm (e.g., Gentner et al., 1993), which led to the conclusion that superficial similarity exerts a stronger effect than structural similarity on retrieval. Given that we took care in ensuring that the target situation did not maintain any degree of structural similarity with the superficially similar base item, the present results are immune to the criticism raised against traditional studies regarding the inadvertent inclusion of structural similarity in the surface matches (see Raynal et al., 2020). On the other hand, as our base analogs consisted of meaningful situations naturally encoded by participants prior to the experimental session, the scarce retrieval of the structural match cannot be attributed to a frugal and mindless processing of the sources.

Even though the events that constitute the plot of popular movies like “Jurassic Park” were available in participants LTM, it could be argued that participants might have been immersed in the constituent episodes of the movies in a local manner, without ever taking the time to compile the key events of the plot in a compact, unified representation like the synopses that appear in press releases and reviews. Even in those cases where such a succinct representation of the complete causal structure was constructed, participants may not have abstracted it at the level of generality that would have been necessary to reliably retrieve other instances of the same underlying logic. In the words of Goldwater, Gentner, LaDue and Libarkin (2021, p. 19): “Even if you have watched Jurassic Park five times, you may not start to see the world in terms of the plot structure of Jurassic Park (at least without active imaginative elaboration)” (For a similar argument, see Raynal et al., 2020).

In order to assess whether the observed supremacy of superficial similarity generalizes to cases where the causal structures of base situations are comparatively easier to compile and generalize, the extra-experimental analogs used in Experiment 2 consisted in highly reproduced public events.

Experiment 2

Method

Participants An initial sample of 388 Argentine residents were invited to volunteer for the study through email, Facebook and WhatsApp. The only requisites for participation were being over 18 years, having completed high school education, and not having taken university-level courses of Psychology. The final sample comprised 98 participants between 18 and 75 years of age ($M = 32.22, SD = 14.63, 67\%$ female) whose scores on the questionnaires presented at the

end of the study demonstrated a faithful memory of at least one of the two public episodes employed as base items.

Materials In order to assess the spontaneous retrieval of the structurally-similar and/or the superficially-similar public events used as base items, participants read the following story:

The city of Ottawa sits on the banks of an important river. About 40 km upriver, a small population of beavers is rapidly decreasing in size. Some time ago, Ottawa's Animal Protection Association has banned the hunting of beavers. The wealthy families from Ottawa have been developing gated neighborhoods on the wetlands of the river. The Animal Protection Association organizes excursions for these families to learn about how beavers live, and how friendly they are. Besides discouraging their hunting, they persuade the families of adopting specimens in order to raise them and breed them, so as to later return the beavers to their natural habitat. The looking of the gated neighborhoods has turned a bit odd, as there are more and more beavers in the gardens of the houses. Recently it came out that members of the Animal Protection Association had been hunting beavers, causing indignation among the families. The hunters apologized, but the Association lost credibility. People kept adopting the beavers, but there are less and less individuals signing for the excursions.

The fictional situation presented above maintains structural similarities with a highly reproduced public episode of 2021, in which the partner of the Argentine President held her birthday party at the presidential residence after the onset of the COVID pandemic, when such encounters were expressly prohibited by the President. This caused indignation in the population, and even though the president apologized, his positive image deteriorated. The structural similarity between the birthday affair and the target analog herein presented has to do with the fact that in both cases (a) authorities impose restrictions, (b) it later comes out that the authorities have themselves violated the restrictions, (c) authorities apologize, and (d) the authorities lose credibility.

At the same time, the target situation presented to participants constitutes a superficial match to another episode of Argentina that went viral during 2021: the proliferation of capybaras in the gated community of Nordelta, a sophisticated neighborhood developed on the wetlands of the Paraná River, a few miles north of Buenos Aires. As these giant rodents have destroyed manicured lawns, bitten dogs and caused traffic accidents, the neighbors are asking the authorities to take action. However, environmentalists argue that capybaras are struggling to recover their once natural habitat, from where they had been excluded. Even though the situation presented to participants shares a few isolated entities and relations with the proliferation of capybaras in Nordelta (e.g., wild rodents in fancy gardens and gated neighborhoods built on wetlands), their causal patterns of events are completely unrelated (e.g., the presence of the rodents is in one of the cases intended due to a decreasing population and in the other resisted due to an overpopulation).

Procedure Participants accepting the invitation to participate in a text comprehension study began by signing an informed consent and stating their age, gender and level of studies. As in Experiment 1, the presentation of the target situation was

preceded by an instruction to read the text very carefully, so as to be able to answer comprehension questions about it. The post-task questionnaires followed the same logic as those of the previous experiment, with the sole difference that instead of alluding to movies, they now alluded to public events (see Table 2). Participants' responses to the multiple-choice questionnaires about the birthday affair and the capybaras invasion determined the assignment of 20 participants to the "birthday" group, 26 participants to the "capybaras" group, and 52 participants to the "birthday + capybaras" group.

Results and Discussion

Among participants who had been exposed to either the superficial match (capybaras) or the structural match (birthday) but not to both, the superficial match was more frequently retrieved than the structural match 50% vs. 0%, $\chi^2(1, 46) = 13.94, p = .0002, \phi = .55$. The central result, however, concerned the responses given by participants who had both a surface match and a structural match available in LTM. Like in the previous contrast, an analysis of the responses given by the birthday + capybaras group revealed that the superficial match was more frequently recalled than the structural match, 38.5% vs. 5.8%, McNemar's $\chi^2(1, 52) = 17, p < .0001, \phi = .57$. This pattern of results therefore generalizes the findings of Experiment 1 to meaningful extra-experimental episodes whose salient structural features are comparatively easier to compile and generalize.

Table 2: Questions used to assess participants' representation of the presidential birthday affair

-
1. This year it became known that Fabiola Yáñez, the President's fiancé, had organized her birthday party at
 - (a) their private flat
 - (b) the presidential residence
 - (c) a bar
 - (d) a rented venue
 2. This meeting
 - (a) was not attended by the President
 - (b) was attended by the President but no other politicians
 - (c) was attended by the President and the vice-president
 - (d) was attended by the President and all the Ministers
 3. In terms of regulations, the birthday meeting
 - (a) did not violate the restrictions arising from the pandemic
 - (b) violated COVID restrictions only slightly
 - (c) only violated the obligation of wearing facial masks
 - (d) did not violate any COVID restrictions at that time
 4. In a later public speech, the President affirmed that
 - (a) the meeting had been illegal
 - (b) the meeting was a mistake
 - (c) the meeting was a healthy relief during a difficult time
 - (d) the meeting was a work-related activity
 5. Those who attended the meeting
 - (a) faced moral condemnation, and were prosecuted by law
 - (b) faced moral condemnation, but were not investigated
 - (c) were criticized by the opposition, but defended by media
 - (d) were legally convicted
-

General Discussion

In previous studies using naturally-acquired base analogs, (Trench & Minervino, 2015; Olguín, Tavernini, Trench, & Minervino, 2022) we have demonstrated that distant analogs are significantly more difficult to retrieve than close analogs. Even though the retrieval disadvantage relative to close analogs might be somewhat inconvenient for some particular tasks (see Olguín, Trench, & Minervino, 2017, for a discussion), the situation would be even more pressing if, given a lack of close analogs in LTM, distant analogs got outcompeted by memory items maintaining only superficial similarity with the target.

Classic studies using a cued-recall paradigm (e.g., Gentner et al., 1993) have found that distant analogs lose the battle against surface matches, wherein objects and isolated first-order relations are shared with the target. However, as demonstrated by Raynal et al. (2020), some proportion of the retrieval of surface matches could conceivably be accounted for by the concurrent presence of structural similarity. As expressed by these authors, the surface matches appear to be literal matches until the outcomes of the stories differ. Such nearly literal similes should not be treated as surface matches since they can often subserve rigorous reasoning. To exemplify, suppose that a friend of yours tells you that she arrived early at an appointment with a busy nutritionist, with the result that she got attended to earlier than originally scheduled. If after hearing this story you recalled one or more episodes wherein having arrived early at some other doctor's appointment did not result in being called in earlier, you would be in a better position to reject your friends' intuitive generalization that arriving earlier than expected would likely lead to leaving earlier from the doctor's office. Hence, the high retrieval of surface matches of the kind employed in some of Gentner et al.'s (1993) sets of materials is neither a conclusive proof of the supremacy of superficial similarity during retrieval, nor an indication about the overall clumsiness of the system, as has sometimes been contended.

But what if distant analogs were outcompeted by LTM items maintaining no trace of structural similarity with the target? Even though Raynal et al.'s (2020) study with a cued-recall paradigm has obtained that purely superficial matches no longer outcompete structural matches, the rather small number of base items included in the learning phase may have allowed participants to serially navigate the learning set in search for a structural match to the target. By concentrating on two types of extra-experimental source analogs, we were able to overcome the serial search objection while attending to Raynal et al.'s, remark about the importance of constructing surface matches maintaining no traces of structural overlap with the target. The results of our two experiments converge in demonstrating that when participants are not hinted about the availability of related content in LTM, the spontaneous recall of meaningful extra-experimental episodes maintaining purely superficial similarity is significantly more frequent than the retrieval of purely structural matches.

Several authors have contended that the superficial bias of our memory systems represents an expectable consequence of the solution given by evolution to the problem of retrieving

structurally related episodes from memory (see e.g., Gentner, 1989). The argument is based on the assumption that the processes responsible for determining whether two situations share an abstract relational pattern are computationally costly, for which carrying out structural matches between the target and all items stored in LTM would be psychologically unrealistic. The evolutionary solution to the problem seems to have consisted in taking advantage of superficial similarities to select a manageable set of candidates for a subsequent structural comparison with the target. As computing surface-level overlap is relatively straightforward, it could conceivably allow a massive search throughout the whole of LTM. Given that superficial features tend to be correlated with more abstract properties (i.e., if two situations share viruses, deceases, countries, vaccines, hospitals, doctors, etc., they likely share the structure of a pandemic), there are chances of finding an analog within the selected set of candidates. On the other hand, as these surface features are relevant for knowledge transfer between analogous cases (i.e., inferences for a pandemic are better when analogizing to another pandemic than when analogizing to a war), the retrieved analogs would likely be among the most useful. However, the cost of this strategy would be to retrieve close analogs and surface matches over distant analogs, thus precluding access to relevant information that could otherwise subserve scientific, artistic, or communicational goals.

In more practical terms, resolving the theoretical question that motivated the present research might bear implications for the design of instructional interventions interested in fighting the problem of inert knowledge. If a portion of this educational problem can in fact be attributed to the surface bias of our memory systems, more efforts should be spent in designing cognitive interventions to neutralize it (e.g., Kurtz & Loewenstein, 2007; Minervino, Olguín, & Trench, 2017; Trench, Tavernini, & Goldstone, 2017, see Trench & Minervino, 2017; 2020 for reviews). But if naturalistic studies are right in claiming that the surface bias is an artifact of crafty experimental procedures, such a rush to develop cognitive protheses would be pursuing a pseudo-problem. Everything seems to indicate that the surface bias is real, and that the quest to develop effective cognitive interventions is heading in the right direction.

Acknowledgments

This work was supported by Grant PIP 11220200103018 from the National Council for Scientific and Technical Research (CONICET) and by Grants PICT 2017-0853, PICT 2019-02542 and PICT 2019-3268 from the National Agency for Scientific and Technological Research (ANPCyT).

References

- Blanchette, I., & Dunbar, K. (2000). How analogies are generated: The roles of structural and superficial similarity. *Memory & Cognition*, 28, 108–124.
- Blanchette, I., & Dunbar, K. (2001). Analogy use in naturalistic settings: The influence of audience, emotion, and goals. *Memory & Cognition*, 29, 730–735.

- Catrambone, R. (2002). The effects of surface and structural feature matches on the access of story analogs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 318–334
- Forbus, K., Gentner, D., & Law, K. (1995). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science*, 19, 141–204.
- Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning*. London: Cambridge University Press.
- Gentner, D., Rattermann, M. J., & Forbus, K. D. (1993). The roles of similarity in transfer: Separating retrievability from inferential soundness. *Cognitive Psychology*, 25, 431–467.
- Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology*, 12, 306–355.
- Goldwater, M. B., Gentner, D., LaDue, N. D., & Libarkin, J. C. (2021). Analogy generation in science experts and novices. *Cognitive Science*, 45, e13036.
- Hesse, M. B. (1966). *Models and analogies in science*. Notre Dame, IN: University of Notre Dame Press.
- Keane, M. T. (1987). On retrieving analogues when solving problems. *Quarterly Journal of Experimental Psychology*, 39, 29–41.
- Kurtz, K., & Loewenstein, J. (2007). Converging on a new role for analogy in problem solving and retrieval: When two problems are better than one. *Memory & Cognition*, 35, 334–341.
- Lakoff, G. & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
- Hofstadter, D. R., & Sander, E. (2013). *Surfaces and essences: Analogy as the fuel and fire of thinking*. New York: Basic Books.
- Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review*, 104, 427–466.
- Minervino, R., Olgún, V., & Trench, M. (2017). Promoting interdomain analogical transfer: When creating a problem helps to solve a problem. *Memory & Cognition*, 45, 221–232.
- O’keefe, D., & Costello, F. (2008). A fast computational model of analogical retrieval (and mapping). In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.
- Olgún, V., Trench, M., & Minervino, R. (2020). The effect of surface similarities on the retrieval of analogous daily-life events. *Memory & Cognition*.
- Olgún, V., Trench, M., & Minervino, R. (2017). Attending to individual recipients’ knowledge when generating persuasive analogies. *Journal of Cognitive Psychology*, 29, 755–768.
- Raynal, L., Clement, E., & Sander, E. (2020). Are superficially dissimilar analogs better retrieved than superficially similar disanalogues? *Acta Psychologica*, 203, 102989.
- Richland, L. E., Holyoak, K. J., & Stigler, J. W. (2004). Analogy use in eighth-grade mathematics classrooms. *Cognition and Instruction*, 22, 37–60.
- Ross, B. H., & Kennedy, P. T. (1990). Generalizing from the use of earlier examples in problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 42–55.
- Spencer, R. M., & Weisberg, R. W. (1986). Context-dependent effects on analogical transfer. *Memory & Cognition*, 14, 442–449.
- Trench, M., & Minervino, R. A. (2015). The role of surface similarity in analogical retrieval: Bridging the gap between the naturalistic and the experimental traditions. *Cognitive Science*, 39, 1292–1319.
- Trench, M., Tavernini, M., & Goldstone, R. L. (2017). Promoting spontaneous analogical transfer by idealizing target representations. In G. Gunzelmann, A. Howes, T. Tenbrink & E. Davelaar (Eds.), *Proceedings of the 39th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society
- Trench, M., & Minervino, R. A. (2017). Cracking the problem of inert knowledge: Portable strategies to access distant analogs from memory. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 66). Academic Press.
- Trench, M., & Minervino, R. A. (2020). *Distant Connections: The Memory Basis of Creative Analogy*. New York: Springer.
- Wharton, C. M., Holyoak, K. J., & Lange, T. E. (1996). Remote analogical reminding. *Memory & Cognition*, 24, 629–643.