Analogical transfer of intentions

Luiza Shahbazyan (l.shahbazyan@gmail.com)
Georgi Petkov (gpetkov@cogs.nbu.bg)
Lilia Gurova (lgurova@nbu.bg)

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

Abstract
Humans are exceptionally good at inferring the intentions behind particular behavior even when the situation is complex or the context is completely new. In this paper we explore the hypothesis that a kind of analogical transfer from past experience to present situations plays an important role in the process of attributing intentions to ambiguous actions. The participants in our experiment were presented with two stories, the latter containing an ambiguous action. They were asked to evaluate how plausible was that the actor in the second story had a particular intention, either positive, or negative, or neutral. We found that the participants rated higher the plausibility of a negative intention when the preceding story was relationally similar and its actor manifested negative intentions. The attribution of intention to the ambiguous action was not different from that in the control condition when the preceding story was dissimilar or perceptually similar, or when its actor manifested positive intentions. These findings suggest that an analogical transfer of intentions does play a role in the attribution of intentions to ambiguous actions but the effect is limited to the attribution of negative intentions.

Keywords: relational similarity; analogical transfer; understanding intentions; hostile attribution bias

Introduction
Imagine that you are working on submitting a joint project with several partners. Just before the deadline, one of the partners, an ex-colleague of yours, calls to apologize that her organization won’t be able to participate due to some legal issues. The withdrawal seriously damages the structure of the proposal and you are not sure whether there will be enough time to negotiate a new partnership or rewrite the framework. And the situation has left you wondering: Did she do it on purpose to sabotage your efforts? Or was it just an unhappy incidence? Or maybe she stepped back in order to protect your project?

We engage in such kind of reasoning on a daily basis and the attributions made have a significant impact on how we encode, interpret, and respond to social events (Baldwin & Baird, 2001; Dodge, 2006). Although we often deal with situations that are novel or ambiguous with regard to the intentions of the actors, it is fascinating that intentional understanding is typically fast, effortless and, to a great extent, reliable. Many researchers subscribe to the view that this is possible due to the generativity of our knowledge system that allows us to infer unavailable aspects of the present by establishing similarities with the past (Penn, Holyoak, & Povellini, 2008; Baldwin, 2002; Heusmann, 1998; Dodge, 2006). Their views diverge, however, when it comes to the mechanisms that enable us to accomplish such inferences. Surprisingly, little empirical evidence has been accumulated so far that might help to resolve the debate. The next section highlights what the main controversy is about.

Controversies over understanding intentions
Within the social information processing paradigm (Crick & Dodge, 1994), researchers have taken for granted that understanding intentions depends on previous knowledge and past experience (Crick & Dodge, 1994; Heusmann, 1998; Dodge, 2006). Heusmann (1998), for example, argued that attributing benign or hostile intentions to others’ actions depends on how elaborated and easily accessed in memory are the hostile and the benign scenarios (schemas) about this type of situation. Such knowledge structures support the inference of missing information that is not available directly from the information input (Burks, Laird, Dodge, Petitit, & Bates, 1999). Consistently, findings from social and developmental research have shown that experience of peer victimization (Yeung & Leadbeater, 2007) and abuse during childhood (Dodge, Bates, & Pettit, 1990) are associated with the tendency to attribute hostile intentions to ambiguous actions, the so called hostile attribution bias (HAB). It was demonstrated as well that negatively-oriented social knowledge predicts over-attrition of hostile intentions (Burks et al., 1999; Dodge, Laird, Lochman, & Zelli, 2002).

According to Penn and his collaborators (Penn & Povinelli, 2007; Penn et al., 2008), however, the mapping of incoming information to perceptually similar’ scenarios cannot explain the sophisticated intentional attribution that humans are capable of. Drawing on the latest findings about how humans and primates understand intentions, they further insisted that “reading” mental states requires an

1 The term “perceptual similarity” is defined here as “similarity between attributes” and it is used as a synonym of “superficial similarity”. Perceptual similarity is usually contrasted with relational similarity, which is defined as “similarity between higher-order relations” and which is used in this paper as a synonym of “structural similarity”.
ability to reinterpret perceptual information in terms of higher-order relational structures coupled with an ability to use this knowledge in drawing inferences about the causal relations between the observable actions and the unobservable mental states that allegedly caused the actions. A similar view, based on their observations of infants and young children, was presented by Baldwin and Baird (2001).

Several studies tried to induce attribution biases by an experimental manipulation of mood by means of perceived social rejection (Dodge and Somberg, 1987), frustration (Orobio de Castro, Slot, Bosch, Koops, and Veerman, 2003), and violent video games (Bushman & Anderson, 2002; Kirsh, 1998). In these studies, however, the effect of mood is not disentangled from other possible effects, including the direct or mediating effects of structural similarity between the present and past situations.

To the best of our knowledge, Graham and Hudley (1994) were the only ones who tried to differentiate effects triggered by negativity from those exerted by the activation of some relational constructs. In their experiment participants were randomly assigned to three conditions and asked to memorize several sentences. The sentences differed across conditions and contained information about a negative intentional outcome of an event, a negative unintentional outcome of an event and a neutral event without causal information. Presumably, the sentences containing negative intentional outcome primed the causal relation between the negative outcome and the hostile intent. In agreement with this expectation, the participants in this condition tended to rate the ambiguous intentions of the actor in another situation as more hostile than the participants in the control condition. However, the authors themselves confessed that even in the intentional condition the priming could be result of the activation of a negative trait, mood or other perceptually similar information. Also, Graham and Hudley did not provide information about was there any significant difference between the control and the unintentional condition. Without such information one cannot say whether the source of the priming influence is perceptual or relational (causal) as stated in their article.

**Insights from research on analogy making**

The role of perceptual and relational similarity in reasoning about causal relations has been studied extensively in the literature on analogy making. It has been proposed that a good analogy reveals the common structure of two situations and thus it makes possible further inferences, which is its main utility (Gentner, 1983). This view is consistent with the conceptualization of the processing mechanism for understanding mental states, articulated by theorists such as Penn and colleagues (2008), Baldwin (2002), and Barnes and Thagard (1997) in the context of empathy.

Because of its recognized generative nature, analogy has often been linked to scientific discovery and problem solving (Holyoak, Gentner, & Kokinov, 2001). Most experimental work so far has been focused on the deliberate and conscious use of analogy. In a typical study people are asked to solve a problem by inferring the solution from a relationally similar situation. For instance, Gick & Holyoak, (1980, 1983) wanted their subjects to solve Duncker’s radiation problem, providing them with the solution of a relationally similar military problem. They did find that significantly more people solved Duncker’s problem when they were given the military situation as a base from where they could transfer the solution.

Research in the past decade has demonstrated, however, that analogical reasoning is not always volitional. It has been shown, for instance, that the perception of a piece of information can be altered by relationally similar information presented beforehand even when participants have not been explicitly instructed to do so (Blanchette & Dunbar, 2002; Kokinov, Feldman & Petkov, 2009). Similarly, Day and Gentner (2007) demonstrated that analogical transfer can occur in situations where there are no specific instruction to make analogies. The participants in their experiments were asked to read texts and answer questions about them. The critical manipulation was that some texts contained two passages that were relationally similar (analogous). The participants’ answers to the questions referring to an ambiguous passage revealed that some inferences from the relationally similar passage were spontaneously made.

However, like in Graham and Hudley’s study (1994), in Day and Gentner’s experiments the initial and the final passages shared both perceptual and relational similarity. This leaves unclear whether the effect will persist without such a strong overlap (Gentner & Smith, 2012). Although these results are not conclusive, they have formed a very promising line of research led by the idea that the analogical transfer based on relational similarity may be the sought inferential process that makes use of existing knowledge to resolve problems such as the attribution of intentions to ambiguous actions. To further explore this idea we designed an experiment in which we tried to disentangle the effects of perceptual and relational (structural) similarity.

**Present study**

The present study aims to test the hypothesis that a relational similarity between two situations would prompt participants to make particular inferences about the unknown intent of the actor in the later situation.

In order to disentangle the effects of negativity, perceptual and relational similarity, we varied 7 base stories in a between subject design to test the effect of two factors: similarity between the base and the target story (relational, perceptual, no similarity) and the ending of the base story (happy-ending, sad-ending). The degrees of relational and perceptual similarity were determined by expert ratings following the definitions of Gentner (1983).

The four conditions of most interest are the ones that involve relational and perceptual similarity. In accordance with the findings of Day & Gentner (2007) and Graham and
Hudley (1994), we expected that “relational similarity, happy-ending” condition will bring to higher ratings of the positive intention than the “no similarity, happy-ending” condition, and, respectively, the “relational similarity, sad-ending” condition will bring to higher ratings of the negative intention than the “no similarity, sad-ending” condition.

It is important to stress that, unlike Day & Gentner’s and Graham and Hudley’s studies, in our study there is no overlap between perceptual and relational similarity in the relationally similar conditions, according to the ratings collected before the study from four experts, familiar with Gentner’s (1983) definition.

Method

Participants

191 young adults (Mean age=20, SD=3.5) participated in the study. Seven of them were excluded because some of their responses were missing or inconsistent.

Materials and procedure

We used a paradigm similar to the one used by Blanchette and Dunbar (2002) and Day and Gentner (2007). Participants were presented first with a story, which was expected to serve as a “base” for an analogy, and later they were presented with another story, a “target”, that missed information about the actors’ intention which could be completed by analogical transfer from the “base”. The participants were instructed to read the first story and remember it for future recall. When the second, ambiguous story was introduced, the participants were asked to rate how plausible it was that the actor had any of three intentions: positive, negative and neutral. In order to prevent the participants from guessing what the objective of the study was, they were instructed that the ratings of the second story were meant for another experiment and that it was included to separate temporally the encoding of the first story from its recall.

Similarly to the study of Graham and Hudley (1994), we devised the target story in a way that allows different interpretations of the actor’s intention and asked participants to rate the plausibility of each of them on a 7-point Likert scale. In order to properly assess the influence of both the sad- and the happy-ending stories, the social situation presented in the target story was neutral, so that the intention was open to three interpretations: positive, negative and neutral.

Table 1 presents a sample of the materials: the target story and the stories used in the sad-ending experimental conditions.

In order to construct relationally similar stories, we followed the convention that a mappable system of interconnected relationships must exist between the two stories (Gentner, 1983). To insure that the relational and perceptual (attributional) similarity do not overlap, the base story in the relational similarity conditions contains animals and animal-like actions that differ from the human characters and human-like actions in the target story, so that “the individual attributes must be left behind in the mapping” (Gentner, 1983, p. 161).

Table 1: Sample stimuli: the target story and the three sad-ending base stories. The happy-ending stories differed only in the last sentence.

| Target story: |
| You met a boy from another school and you became friends despite he was older than you. You were hanging out together and you slept over his family’s house several times. You also liked his friends, though they teased you sometimes for being a baby because your parents gave you a curfew. To prove them wrong, sometimes you stayed later than you were allowed. One evening your parents got very angry and forbade you to meet your friend. You went to your room, explained the situation to him over Facebook and he advised you to stay at his place for some time in order to stress your parents, so that they could change their attitudes towards you. |

| Base 1: Relational similarity, sad-ending |
| Once upon a time people bred bear cubs in pens and sheared them like sheep to knit warm sweaters. The pens were guarded by dogs, so that the cubs didn’t get away or got eaten by wolves. One day the wolves sent messengers to go to the cubs to make peace. The wolves said that they really empathize with the cubs that they are wild animals who have to live in captivity, guarded by dogs. The wolves promised they won’t hurt the cubs and offered an escape plan in which several wolves would distract the dogs, while the cubs escaped from the back of the pen. The cubs agreed immediately, because they were tired of being kept against their will. In the morning, the cubs waited for the dogs to start chasing several wolves, left from the back door of the pen and started running. Immediately, the rest of the wolves attacked and killed them all. |

| Base 2: Perceptual similarity, sad-ending |
| Ivo has a birthday and he is wondering how to celebrate it. He wants to gather all his friends, but he knows that his parents would never give him that much money. He shared the concerns with a few friends and one of them invited him to organize a party at his grandfather’s summer house in a nearby village, where the guests can sleep over. Ivo needs only to buy some food and drinks. He agreed it means that he can invite the entire class. The next day, an invitation has been sent out on Facebook and all classmates confirmed. The next day everybody was supposed to come to the bus station, so that they can travel together to the summer house, but only half of the confirmed guests arrived. Ivo felt very miserable, but his friends convinced him to celebrate after all. Several days after that Ivo realizes that some of his classmates deliberately confirmed although they knew that they would not come in order to ruin his birthday party. |

| Base 3: No similarity, sad-ending |
| Due to heavy rainfall a landslide was activated in a mountainous road. Tons of mud poured on the road blocking traffic. Special teams were immediately sent to the place of the accident to clean and strengthen the road. Some time before dawn one of the workers noticed that a redish metal object is buried under the slurry of rock and mud. Together with several other workers, he tried to dig out the mysterious object, but the task was very dangerous due to the instability of the slope. Soon, it became clear that the object is a car. After several hours of digging the workers found out that the driver of the car was dead. He was identified as an 82 years old man from the nearby village. The autopsy showed that he died from heart attack, probably shocked by the landslide. |

In both relational and perceptual similarity conditions the stories contained a causal relation between the intention of some characters of the story and their actions towards the main character. However, in the perceptual similarity conditions the base stories do not share a system of interconnected relationships with the target story, unlike in the relational similarity conditions, but only common attributes. The stimuli were rated by members of our lab,
who are aware of Gentner’s (1983) definition of similarity but are naïve as to what is the objective of the study. According to the averaged ratings of 4 experts on a 7-point Likert scale, the relational similarity in the relational similarity condition is 5, while the perceptual similarity is 1. Consistently, the perceptual similarity in the perceptual similarity condition is 6.25, while the relational similarity is 3. No intentional information was included in the base stories in the no similarity conditions, including the control condition.

The gender of the characters of the stories, when applicable, was congruent with the gender of the participant. For example, the word “boy” in the target story (Table 1) was changed with “girl” for the females. All seven versions were distributed randomly among participants. Participants read one of 7 different base stories but always viewed one and the same story as a target. The stories were presented in a 3-page paper-and-pencil form. The first page contained the instruction for reading and memorizing the base story followed by the story itself. On the second page there were instructions for reading the story and rating the plausibility of any of the intentions followed by the story itself and the three intentions to rate on a 7-point Likert scale from 1 – “I would never consider it” to 7 – “I would definitely consider it”. The participants were explicitly asked to imagine that they were the main character in the story. On the third page there was instruction for participants to write down their recollection of the base story.

Participants were instructed to work on the tasks according to the timing provided by the researcher. They were presented with a cover story and invited to start with the first task. The time limit for the reading of the base story was 2 minutes. Then, the researcher instructed participants to proceed to page 2 to rate the intentions of the character in the target story. After another 2 minutes, participants were instructed to proceed to page 3 in order to write down what they remember from the first story at their own pace. After the end of the study, the participants were debriefed and invited to ask questions.

Results

There were 7 groups and 3 dependent variables: mean rating of the positive intention, mean rating of the negative intention and mean rating of the neutral intention. All means are presented in Table 2.

First, we performed three ANOVA analyses for participant’s rating of each of the three intentions: positive, negative and neutral. All three ANOVA analyses used two independent variables – the type of similarity between the base and the target story (relational, perceptual, no similarity), and the type of ending of the base story (happy-ending, sad-ending). The control group was excluded from this analysis because it had a neutral (neither happy nor sad) ending. There were no main effects or interactions observed in the pattern of ratings of the positive and neutral intentions. There was significant interaction between similarity and ending in the ratings of the negative intention, F(2, 159) = 3.396, p = 0.036.

Second, we performed separate t-tests, comparing each of the groups with the control one. There were significant differences only for the relational, sad-ending condition: t(50) = -3.075, p = 0.003 for the judgments of the positive outcome and t(50) = 2.531, p = 0.015 for the judgments of the negative outcome.

In addition, we analyzed only the ratings of the negative intention if preceded by sad-ending story (relational, perceptual, no similarity). The one-way ANOVA shows significant difference among groups: F (2, 77) = 3.876, p = 0.025. According to the post-hoc test, the difference was between the relational and the no similarity group (p = 0.007), whereas the perceptual similarity group did not differ from the other two (see Figure 1).

These results support the hypothesis that reading a sad-ending base story that is relationally similar to the target story will produce significantly higher rating of the plausibility for a negative intention than reading a sad-ending but dissimilar base story or a neutral story. Consistently, it produced significantly lower rating for a positive intention than reading a neutral story. However, we failed to replicate the previous findings suggesting that perceptually similar information and the negativity of the story on its own could induce attribution bias (Dodge and Somberg, 1987; Orobio de Castro, et. al., 2003; Bushman & Anderson; Kirsh, 1998).

Discussion

A number of theorists suggested that intentional understanding must involve some kind of inductive mechanism which draws on the similarity between new situations and past experiences based on higher-order, nonobvious relations. In accordance with this approach, we demonstrated that an analogical transfer from a sad-ending story resulted in higher ratings of the negative intentions and lower ratings of the positive intentions attributed to an ambiguous action. It is important to stress that this result could not be reduced to the effect of negativity of the ending
or perceptual similarity because no such effects were observed in the other conditions (perceptual similarity, no similarity and control).

Our results contribute also to current analogy research by showing that analogical transfer based on relational similarity plays a role in understanding new situations, even without explicit instructions. Given the very low rate of spontaneous use of analogies in studies using tasks like Duncker’s radiation problem, results like these are a bit surprising. For instance, Day and Gentner (2007) suggested that surface (or perceptual) similarity is, probably, necessary for a nondeliberative analogical transfer to take place. Our findings show that maybe this is not always the case. On the contrary, it seems that only relational similarity has impact on the analogical transfer that prompt attributing negative intentions. We should admit, however, that the reported here results are not entirely conclusive because no significant difference was established between the effects of the “perceptual similarity, sad-ending” condition and the “relational similarity, sad-ending” condition.

The fact that our experiment failed to replicate the effect of perceptual similarity and negativity (sad-ending) needs to be further explored. One explanation for this might be that the effects obtained by Day & Gentner (2007) and Graham and Hudley (1994) are in fact produced by the relational similarity. Thus, alternative explanations of their findings in terms of negativity or perceptual similarity can be ruled out.

This explanation, however, is not applicable to other pieces of research where a hostile attribution bias was successfully induced by exposing participants to negative situation. It could be that a single presentation of a negative material is much weaker manipulation than the ones used in the mentioned studies of hostile attribution bias. In the study of Dodge and Somberg (1987), for example, negativity was induced by explicit peer rejection, which might be much more manipulative for the participants than the hypothetical low-arousal situations which we used. In a similar way, Bushman and Anderson (2002) and Kirsch (1998) successfully induced over-attribute of negative intentions by giving their participants to play violent video games, all of which (except one) were rated appropriate for people aged 18 and over due to the extreme violence (strong language) as rated by PEGI (http://www.pegi.info).

The fact that the manipulation of negativity, which we used, is rather weak might be able also to explain why the effect of relational similarity was found only on the ratings of the negative intentions. On the other hand, however, nobody (to the best of our knowledge) has demonstrated so far experimental manipulation of positive attribution bias. Having that in mind, we may guess that the significant effects of analogical transfer of negative intentions which we observed is a kind of a joint effect of the general preference towards negative interpretations and the experimental manipulation (the sad-ending relatively similar base story) which we made.

The demonstrated asymmetry between the effects exerted by the happy-ending and the sad-ending story bases is in line with the findings that negative emotions, anxiety in particular, enhance relational encoding (Hristova & Kokinov, 2011). It could be that in the sad-ending condition, by asking participants to imagine that they are the main character in the story, we had increased their level of negative arousal. As a result, they had become more sensitive to the relations in the base story and thus encoded them better. The latter allowed them to see the relational similarity and, drawing on it, to make inferences about the target story. Further analysis of the memory task could shed light on this issue.

The results of our study which we reported in this paper do not allow choosing between the different explanations which we sketched here. Further research is needed to assess in a more systematic way the influence of ending, whether happy or sad, of the base stories by controlling for the level of arousal and relational encoding.

Although preliminary, the results of the present study extend our current understanding of intentions in several ways. First, it was demonstrated that no overlap between perceptual and relational similarity is necessary for people to spontaneously use analogical transfer to infer unknown properties of the situation. Second, it was demonstrated that structural-analogical transfer interplays with one of the core social cognitive functions: understanding other’s intentions. In a way, we provided the first empirical evidence supporting the hypothesis that understanding intentional mental states may involve an inferential process based on transfer of higher-order relational information.

Several limitations of the study are worth mentioning. First, it could not be convincingly stated that the obtained effects are not stimulus specific. Further replications with additional sets of stimuli have to be made to ensure that the results are reliably established. Besides, additional controls

<table>
<thead>
<tr>
<th>Condition</th>
<th>Relational, happy-ending</th>
<th>Relational, sad-ending</th>
<th>Perceptual, happy-ending</th>
<th>Perceptual, sad-ending</th>
<th>No similarity, happy-ending</th>
<th>No similarity, sad-ending</th>
<th>Control (no similarity, neutral ending)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive intention</td>
<td>2.66 (1.56)</td>
<td><em><em>1.88</em> (1.34)</em>*</td>
<td>2.66 (1.88)</td>
<td>3.00 (1.85)</td>
<td>2.78 (1.69)</td>
<td>2.69 (1.59)</td>
<td>3.27 (1.87)</td>
</tr>
<tr>
<td>Negative intention</td>
<td>4.66 (1.97)</td>
<td><em><em>5.81</em> (1.44)</em>*</td>
<td>4.41 (2.15)</td>
<td>4.93 (2.09)</td>
<td>5.11 (1.78)</td>
<td>4.35 (2.10)</td>
<td>4.62 (1.97)</td>
</tr>
<tr>
<td>Neutral intention</td>
<td>3.66 (1.95)</td>
<td>3.96 (2.13)</td>
<td>4.03 (2.53)</td>
<td>3.36 (1.77)</td>
<td>3.37 (1.98)</td>
<td>3.54 (2.04)</td>
<td>3.38 (1.96)</td>
</tr>
</tbody>
</table>

Table 2. Mean ratings (standard deviations) of intentions by type and condition.
should be introduced in order to reveal the source of the asymmetry between the ratings of the positive and the negative intentions. More sensitive measures such as RTs could be employed to further explore the pattern of influence of perceptual and relational similarity on the understanding intentions. We believe these are important issues to advance and important tracks for further research to follow.

References


