The dark side of cognitive illusions: When an illusory belief interferes with the acquisition of evidence-based knowledge

Ion Yarritu¹, Helena Matute¹* and David Luque²,³

¹Department of Experimental Psychology, Deusto University, Bilbao, Spain
²Biomedical Research Institute (IBIMA), University of Malaga, Spain
³School of Psychology, UNSW, Sydney, Australia

Cognitive illusions are often associated with mental health and well-being. However, they are not without risk. This research shows they can interfere with the acquisition of evidence-based knowledge. During the first phase of the experiment, one group of participants was induced to develop a strong illusion that a placebo medicine was effective to treat a fictitious disease, whereas another group was induced to develop a weak illusion. Then, in Phase 2, both groups observed fictitious patients who always took the bogus treatment simultaneously with a second treatment which was effective. Our results showed that the group who developed the strong illusion about the effectiveness of the bogus treatment during Phase 1 had more difficulties in learning during Phase 2 that the added treatment was effective.

Human cognition has shown to be prone to a biased interpretation of reality. People tend to believe falsely that they are better than others (Brown, 1986; Pronin, Lin, & Ross, 2002), that their own skills can determine their success in a purely chance task (Langer, 1975), or that certain bogus treatments they follow can miraculously cure their diseases (Matute, Yarritu, & Vadillo, 2011). These false beliefs, typically known as cognitive illusions, have often been related in the psychological literature with mental health and well-being (Lefcourt, 1973; Taylor, 1989; Taylor & Brown, 1988). However, do cognitive illusions have beneficial consequences in all cases? Current discussion in the literature suggests that whereas biases and illusions can often contribute to adaptive adjustment, this is not always the case (see McKay & Dennett, 2009 for an extensive review).

One psychological approach states that cognitive illusions are an adaptive mechanism, ensuring the correct fitness of the person to the environment (Taylor & Brown, 1988). From this perspective, the cognitive system has evolved to interpret the world unrealistically, in a manner that assures the protection of the self. In this framework, illusions related to the perception of relationships between events, such as illusory correlations (Chapman & Chapman, 1969), illusions of control (Alloy & Abramson, 1979; Langer, 1975), or causal attributional biases (Kelley, 1972), are typically assumed to
have an important role in psychological well-being (Taylor & Brown, 1988). It has been argued that instead of interpreting the environmental information rationally, people tend to adjust the environmental data to their prior conceptualization of the world in a manner that is self-serving (Fiske & Taylor, 1984; Lefcourt, 1973; Nisbett & Ross, 1980; Zuckerman, 1979). For instance, it has been found that the illusion of control, a bias by which people tend to overestimate their own control over uncontrollable outcomes (Langer, 1975), works differently as a function of mood, which has sometimes been interpreted as supporting its role as a self-esteem protection mechanism. Whereas non-depressive people view themselves as controlling outcomes which are actually uncontrollable (i.e., illusion of control), depressive people detect the absence of any relationship between their actions and the desired outcomes. This has been called depressive realism (Alloy & Abramson, 1979; Blanco, Matute, & Vadillo, 2012; Msetfi, Murphy, Simpson, & Kornbrot, 2005). Given that the perception of uncontrollability is related to helplessness and depression (Abramson, Seligman, & Teasdale, 1978), some researchers have suggested that either depressed people are depressed because they do not show an illusion of control, or they do not develop the illusion because they are depressed (Alloy & Clements, 1992). In either case, this is an example of how the illusion of control could be related to well-being under this framework (but see Blanco et al., 2012; Msetfi et al., 2005 for more neutral interpretations of this illusion).

A rather different approach suggests that cognitive illusions are just the by-products of a cognitive system which is responsible for extracting knowledge about the world (Beck & Forstmeier, 2007; Haselton & Nettle, 2006; Matute et al., 2011; Tversky & Kahneman, 1974). The discussion revolves nowadays around the benefits and costs of establishing false beliefs (Haselton & Nettle, 2006). From this point of view, cognitive illusions are not beneficial per se. Instead, they would be the necessary cost to be assumed by an overloaded cognitive system that tries to make sense of a vast amount of information (Tversky & Kahneman, 1974). The results of this assumable cost range from superstitious behaviour, magical thinking, or pseudoscientific beliefs (Matute, 1996; Matute et al., 2011; Ono, 1987; Vyse, 1997) to prejudice, stereotyped judgements, and extremism (Hamilton & Gifford, 1976; Lilienfeld, Ammirati, & Landfield, 2009). The previously mentioned self-serving illusions (Taylor & Brown, 1988) would be interpreted as part of this cost under this view.

Therefore, while keeping in mind that the cognitive illusions can eventually lead to benefits related with psychological well-being, there are also cases in which their collateral costs can lead to serious negative consequences. Take as an example, the person who develops the false belief that a pseudoscientific (i.e., innocuous, at best) treatment produces the recovery from a disease from which he or she is suffering. Believing that the pseudoscientific treatment is effective, that person could underestimate the effectiveness of a medical treatment which actually works. This bias could lead the person to reject the really effective treatment and, consequently, suffer the implications derived from this action. Or, in another example, if a person believes that a certain minority group has higher rates of delinquency, how could we convince that person, at the light of evidence, that his/her belief is not true? The two scenarios drawn here are examples that show that a false belief could, under certain conditions, interfere with the establishment of grounded, evidence-based, knowledge.

Despite the theoretical and practical relevance of this problem, there are, to our knowledge, very few studies focusing on how illusory beliefs affect the acquisition of evidence-based knowledge. One of the very few studies we are aware of is that of Chapman and Chapman (1969). They found that illusory correlations in the interpretation
of projective tests could blind psychologists to the presence of valid correlations between symptoms. However, it is not clear in their study how the illusions were developed, nor the mechanism by which their occurrence could blind the detection of real correlations. While the applied nature of their study was certainly commendable, it implied that several aspects outside of experimental control, such as previous knowledge, credibility of the source from which the illusion was acquired, strength of the belief, or years psychologist had maintained the illusory belief, could, at least in principle, be affecting the results. Because their goal was highly applied, Chapman and Chapman did not create the different experimental conditions and manipulations over these illusory correlations, as they only selected the most frequent erroneous interpretations of a projective test. The main goal of the present work is to explore, in an experimental setting, the potential interference that illusory beliefs might exert over the subsequent learning of new evidence-based knowledge, and to propose a broad mechanism by which this could occur.

Cognitive illusions usually involve beliefs about causal relationships between events that are, in fact, unrelated (i.e., illusion of causality). For instance, the illusion of control involves the belief that our own action (the potential cause) produces the occurrence of the desired goal (the outcome). The experimental literature on causal learning is a fruitful framework for studying these cognitive illusions (Matute et al., 2011). Many causal learning experiments have shown that learning about the relationship between a cause and an effect influences the subsequent learning of another cause that is paired with the same outcome. The family of learning phenomena known as cue interaction represents the way by which these effects occur. When two potential causes, A and B, are presented simultaneously and paired to an outcome, they compete for establishing a causal association with that outcome. In these cases, the existence of previous experience or previous knowledge about the relationship of one of the causes and the outcome determines what can be learned about the second cause. For instance, the learner may believe that one of the potential causes produces the outcome or, on the contrary, the learner may believe that one of the causes prevents the occurrence of the outcome. In both cases, this previous belief about one of the causes, say A, will affect what can be learned about the other cause, B, when both causes are presented together and followed by the outcome. In the first case, when the previous belief is that A is the cause of the outcome, the detection of a causal relationship between the second cause B and the outcome will be impaired (i.e., this particular case of cue interaction is generally known as the blocking effect; Kamin, 1968). In the second case, when the previous belief is that A prevents the outcome from occurring, the detection of a causal relationship between the second cause B and the outcome will be facilitated (i.e., this particular case of cue interaction is generally known as superconditioning; Rescorla, 1971). Many cue interaction experiments, both with animals and humans, show that learning about the relationship between a potential cause and an outcome can result altered when the potential cause is presented in compound with another potential cause that has been previously associated either with the outcome or its absence (Aitken, Larkin, & Dickinson, 2000; Arcediano, Matute, Escobar, & Miller, 2005; Dickinson, Shanks, & Evenden, 1984; Kamin, 1968; Luque, Flores, & Vadillo, 2013; Luque & Vadillo, 2011; Morís, Cobos, Luque, & López, 2014; Rescorla, 1971; Shanks, 1985).

Therefore, given that previous causal knowledge can interfere with the learning of new causal knowledge, and given that previous knowledge could in principle be illusory, a question of interest is whether the development of cognitive illusions could interfere with the development of new and evidence-based causal knowledge. To answer this
question, we designed the current experiment, using a standard contingency learning task (Wasserman, 1990). In our experiment, participants learned about the effectiveness of some medicines through observation of fictitious patients: The fictitious patients either took a medicine or not, and they either recovered from the crises produced by a fictitious disease or not (Fletcher et al., 2001; Matute, Arcediano, & Miller, 1996). The experiment was divided into two learning phases. In the first phase, participants were exposed to information that should induce the illusion that a medicine (Medicine A) that had no real effect on the patients’ recovery was nevertheless effective. In this phase, two groups of participants differed in the information they received. For one group, the illusion was induced to be high and for the other was induced to be low (see Method). In the second phase, the ineffective medicine used in the first phase, Medicine A, was always presented in compound with a new medicine (Medicine B), which actually did have a curative effect over the patients’ disease. The question was whether the acquisition of an illusory causal relationship between the (ineffective) Cause A and the outcome during Phase 1 would interfere with subsequent learning about the causal relationship between the potential (and in this case, actually effective) Cause B and the same outcome that was presented during Phase 2. We expected that the different degree of illusion about Medicine A induced in both groups during Phase 1 would lead participants of the two groups to assess the effectiveness of Medicine B (i.e., the effective one) differently at the end of Phase 2. More specifically, we expected that the group for which we induced higher illusions about the effectiveness of Medicine A should show greater difficulties than the other group in detecting that Medicine B was actually effective.

Method

Participants and apparatus
We recruited 147 university students, who participated in the experiment in exchange for academic credit. Any student in the Psychology of Learning class who expressed their willingness to participate was allowed to do so. Participants were randomly assigned to each of the two groups, resulting in a total of 73 participants in the high-illusion group and 74 participants in the low-illusion group. Participants performed the task on personal computers. The program was implemented as an HTML document dynamically modified with JavaScript.

Ethics statement
The data that the participants provided were anonymous and unidentifiable, the stimuli and materials were harmless and emotionally neutral, the goal of the study was transparent, and the task involved no deception. The participants were informed before the session that their data would not be identifiable and that they would be allowed to terminate the study by closing the task program window at any time without penalty, if they wished so. In addition, right after the study finished, a screen asked for the participants’ permission to send us the data they had just generated. Only the data from those participants who granted their permission by clicking a button labelled ‘Send data’ were stored and used herein. Those participants not willing to submit their responses had the option of clicking a button labelled ‘Cancel’, which immediately deleted the data. The procedure was approved by the university ethics committee.
Procedure and design
The task was an adaptation of the allergy task, which has been widely used in causal learning studies. This task has proven to be sensitive to the effect of the illusion of causality (Matute et al., 2011). Participants were prompted to imagine being a medical doctor, who specialized on a rare disease called ‘Lindsay syndrome’. They were then told that there existed some new medicines that could cure the crises caused by the disease. Their mission was to find out whether these medicines were effective. We used two fictitious names for the medicines, ‘Batatrim’ and ‘Aubina’. These two names were counterbalanced so that for some participants, the first trained medicine was Batatrim and for the other participants was Aubina. The experiment comprised two training phases, each containing 100 trials. In each trial, participants could first see whether a fictitious patient had taken the medicines or not (potential cause) and then observed whether the patient recovered from the crises (outcome). In trials in which the medicine was taken, participants saw a picture of the medicine (a picture of Batatrim or Aubina in the first phase and a picture of the two medicines together in the second phase) and the sentence ‘The patient has taken’ and the name of the medicine (or medicines). When the medicine was not taken, participants saw the sentence ‘The patient has not taken medication’ and no picture was presented. Immediately underneath the information about the medicine, they could read the question ‘Do you think the patient will recover from the crisis?’ This prediction question was used to maintain the participants’ attention and to make sure they were reading the screen. They could answer by clicking on one of two buttons (yes/no). Once they gave their responses, a third (lower) panel showed the information about whether the patient had recovered or not. In trials in which the patient had recovered, participants saw a picture of the patient recovered and the sentence ‘The patient has recovered from the crisis’. When the patient had not recovered, participants saw a picture of the patient affected by the crisis and the sentence ‘The patient has not recovered from the crisis’ (see Figure 1 for an example of how these trials were presented).

Illusions of causality are found to be strongly affected by the frequency with which the potential cause and the outcome occur. When the outcome occurs with a high probability, the illusion is stronger (Allan & Jenkins, 1983; Alloy & Abramson, 1979; Blanco, Matute, & Vadillo, 2013; Matute, 1995; Shanks & Dickinson, 1987). In addition, when the probability of the potential cause is high, the illusion will also be stronger (Blanco et al., 2013; Hannah & Beneteau, 2009; Matute, 1996; Matute et al., 2011; Perales & Shanks, 2007). These two factors are often known as density biases (i.e., cue density and outcome density), and they play a crucial role in the development of false beliefs about causal relationship (Allan & Jenkins, 1983; Hannah & Beneteau, 2009; Matute, 1995, 1996; Matute et al., 2011; Yarritu, Matute, & Vadillo, 2014). The illusion is particularly strong when both the cause and the outcome occur frequently (Blanco et al., 2013). To manipulate the degree of the illusion of causality developed by our participants, we used a high probability of the outcome in all cases and manipulated between groups the frequency of occurrence of the potential cause during Phase 1.

Table 1 shows a summary of the experimental design. During Phase 1, the potential cause was a single medicine (A) which had no causal relationship with the outcome. That is, the sequence of potential cause–outcome pairings was programmed in such a way that the outcome occurred with the same probability in the presence and in the absence of the potential cause. The probability of occurrence of the outcome was high (.70) in both groups because, as described above, this is known to produce illusions. In
In this phase, however, we manipulated the probability of occurrence of the potential cause, so that it was different for the two groups of participants. For the high-illusion group, this probability was .80, whereas for the low-illusion group, it was .20.

**Table 1.** Design summary

<table>
<thead>
<tr>
<th>Group</th>
<th>Phase 1</th>
<th>Phase 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$p(A)$</td>
<td>$p(O</td>
</tr>
<tr>
<td>High Illusion</td>
<td>.8</td>
<td>.7</td>
</tr>
<tr>
<td>Low Illusion</td>
<td>.2</td>
<td>.7</td>
</tr>
</tbody>
</table>

*Note.* $A$ and $B$ are fictitious medicines. $O$ (outcome) is recovery from the crises produced by Lindsay syndrome. Med = Medication.

In this phase, however, we manipulated the probability of occurrence of the potential cause, so that it was different for the two groups of participants. For the high-illusion group, this probability was .80, whereas for the low-illusion group, it was .20.
mentioned earlier, when the probability of the outcome is high, a high probability of the potential cause leads to a stronger illusion of causality than does a low probability (Blanco et al., 2013; Matute et al., 2011). After completing all 100 training trials of this phase, participants were presented with the following question: To what extent do you think that Batatrim (or Aubina) was effective in healing the crises of the patients you have seen? The answers were given by clicking on a 0–100 scale, anchored at 0 (definitely NOT) and 100 (definitely YES). This judgement question was introduced at the end of Phase 1 to make sure that our procedure led to a stronger illusion in the high-illusion group than in the low-illusion group.

Once this question was answered, participants were told that they would see the remaining patients and the second training phase began. This phase was identical for both groups. In this phase, a new medicine (B) always appeared in compound with the medicine trained in the previous phase (i.e., A + B). This new medicine was Aubina if Batatrim was trained in the previous phase or Batatrim if it was Aubina which was trained previously. The compound was presented in half of the training trials, that is half of the fictitious patients in this phase took both medicines simultaneously, whereas the other half took none. The probability of the outcome in the presence of the compounded medication was higher (.90) than in its absence (.70). That is, the new medicine had a positive effect on the healing of the crises of Lindsay syndrome because its intake implied an increment in the probability of recovery. After completing all 100 training trials of this second phase, participants were asked to emit their judgement about the new medicine (B), which was our target-dependent variable. This judgement was worded in the same way as the previous one. The participants’ answers were also given in the same way as in the previous phase.

Results

We first made sure that our manipulation was effective in promoting a stronger illusion in the high-illusion group than in the low-illusion group by the end of Phase 1. Means (and standard errors of the means) of the effectiveness judgement for Medicine A at the end of Phase 1 were 65.42 (2.00) for the high-illusion group and 47.95 (2.58) low-illusion group. To discard the potential effect of the counterbalancing of the names of the medicines, we conducted a 2 (group) × 2 (counterbalancing) ANOVA on the judgements of Medicine A. Neither the main effect of counterbalancing nor the interaction between these two factors was significant (lower p = .13). As expected, however, the main effect of group resulted significant, $F(1, 143) = 28.17, p < .001, \eta^2_p = 0.16$. Thus, our manipulation was successful.

The critical results of this experiment are the judgements about Medicine B at the end of Phase 2. Means (and standard errors of the means) of the judgement of effectiveness for Medicine B at the end of Phase 2 were 67.42 (2.49) for the high-illusion group and 74.36 (2.31) the low-illusion group. To discard the potential effect of our counterbalancing procedure, we conducted a 2 (group) × 2 (counterbalancing) ANOVA on the judgements of Medicine B. Neither the main effect of counterbalancing nor the interaction between counterbalancing and group was significant (lower $p = .37$). As expected, however, the main effect of group was significant, $F(1, 143) = 4.05, p < .05, \eta^2_p = 0.03$, showing that the high-illusion group gave a lower judgement for Medicine B (which actually was effective) in Phase 2.
Discussion

Holding the illusory belief that a bogus treatment is efficient leads to a stronger reticence to accept that an evidence-based treatment works better than the bogus one. Our results show that the group who was induced to develop a strong illusory belief about the effectiveness of an inefficient treatment (Medicine A) judged the actually effective treatment (Medicine B) to be less effective, as compared to the group who was induced to develop a weaker illusion. In the experiment presented here, the second training phase, in which the two medicines were trained in compound, was identical for both groups. Therefore, the differences in how both groups assessed the effectiveness of Medicine B during the second phase must have resulted from the manipulation conducted during the first phase, in which only Medicine A was presented. Note, also, that the evidence presented to participants in both groups during the first phase could only support the objective conclusion that Medicine A was totally ineffective because the probability that the fictitious patients recovered was the same regardless of whether they took Medicine A or not. Thus, if the participant’s judgements would have been accurate during Phase 1, then participants in both groups should have learned exactly the same about Medicine A, that is, that Medicine A was completely ineffective. This learning should have equally affected their learning about the effectiveness of Medicine B in both groups during Phase 2. However, the participants’ judgements show that this was not the case.

The unique factor that can explain the differences in the judgements for Medicine B between both groups at the end of Phase 2 is the difference in the degree of the illusion developed in the first phase (as shown by the judgements about Medicine A). Thus, as expected, previous training on an illusory belief about Medicine A exerted an influence over the establishment of a true evidence-based belief concerning the effectiveness of Medicine B. This influence is an example of a phenomenon that we already described above, cue interaction. In the second phase of the present experiment, the two potential causes, Medicine A and Medicine B, competed for establishing the causal relationship with the outcome, that is, with the patients’ recovery. However, the two groups of participants had received different exposure to one of the causes (Medicine A) in the first phase, a manipulation that is known to induce a stronger illusion of causality. Thus, one of the groups had learned illusorily that Medicine A and the patient’s recovery were causally related, whereas for the other group, this illusion was significantly weaker. This differential exposure to Medicine A during the first phase led to differences between the two groups in their effectiveness judgements of Medicine B. Given the experimental design, we are not in a position to discriminate whether the critical differences in the way the two groups judged the effectiveness of Medicine B at the end of Phase 2 should be attributed to the high illusion that Medicine A was effective reducing the judgement about Medicine B in the high-illusion group (as in the blocking effect; see above) or to the lower (i.e., more accurate) perceived effectiveness of Medicine A in the low-illusion group producing an overestimation in the assessment of Medicine B in that group (as in the superconditioning effect; see above). Quite possible, the two effects contributed to the observed differences, as is often the case in the cue interaction literature.

Nevertheless, what is clear given the present results is that the group who developed a high illusion about the ineffective Medicine A tended to assess the effective Medicine B as less effective than the group who developed a weaker illusion. Similar cue interaction effects have been clearly established in other causal learning research (Aitken et al., 2000; Arcediano et al., 2005; Dickinson et al., 1984; Luque & Vadillo, 2011; Luque et al., 2013; Morís et al., 2014; Shanks, 1985). Moreover, cue interaction effects are beautifully
predicted by current theories of learning (Rescorla & Wagner, 1972). Here, we show that illusions, and not only evidence-based knowledge, can compete with the acquisition of new knowledge and can produce cue interaction effects like those already known to occur in response to previous learning.

Last but not least, these results contribute to the currently open debate about the potential adaptive value of cognitive illusions. More specifically, these results are consistent with the nowadays growing opinion that cognitive illusions and biases are not essentially adaptive or non-adaptive per se, but rather, that they should be considered in the context in which they appear and under the light of the mechanisms that generate them (McKay & Dennett, 2009). In the case of the experiment presented herein, learning about an evidence-based treatment was impaired in the group who developed a stronger illusion as compared to the group who developed a weaker illusion. The consequences of this particular cognitive illusion cannot be regarded as adaptive. The example presented above can clarify this point: If a person believing that a pseudoscientific treatment works misperceives the effectiveness of an evidence-based treatment and, following this misperception the person rejects the effective treatment, that person can undergo terrible consequences, sometimes even death. A similar example is taking place today at most Western countries when people reject vaccination on the argument that they do not promote health. If someone lives in a city where everyone else has undergone vaccination, then this person will not suffer from certain diseases. The problem will be that the illusory attribution of causality (i.e., I am fine because this disease is not a real risk) will compete with the acquisition of evidence-based knowledge on the effectiveness of vaccination programmes. The present study shows how this biased thinking can develop. We are not saying that all cognitive illusions do necessarily compete with newer evidence-based knowledge. However, illusory beliefs have often demonstrated an atypical persistence, in spite of new evidence contradicting them (Chapman & Chapman, 1969; Nisbett & Ross, 1980). Revisiting those persistent and often serious cases, at the light of the results of the present study, could possibly prove fruitful.

Acknowledgements

We would like to thank Rafael Alonso-Bardón, Itsaso Barberia, Fernando Blanco, Pedro L. Cobos, Carmelo P. Cubillas, Amanda Flores, Francisco J. López, Joaquín Morís, Nerea Ortega-Castro, and Miguel A. Vadillo for illuminating discussions. We also would like to thank Scott Lilienfeld and an anonymous reviewer for helpful comments and suggestions on a previous draft. Funding support for this research was provided by Grant PSI2011-26965 from Dirección General de Investigación of the Spanish Government and Grant IT363-10 from Departamento de Educación, Universidades e Investigación of the Basque Government, both granted to Helena Matute.

References


Received 12 February 2014; revised version received 10 December 2014