The automatic nature of motivated belief updating

ANDREAS KAPPES*

Department of Psychology, City, University of London, London, UK

TALI SHAROT

Department of Experimental Psychology, University College London, London, UK

Abstract: People's risk estimates often do not align with the evidence available to them. In particular, people tend to discount bad news (such as evidence suggesting their risk of being involved in a car accident is higher than they thought) as compared to good news (evidence suggesting it is lower) – this is known as the belief update bias. It has been assumed that individuals use motivated reasoning to rationalise away unwanted evidence (e.g., "I am a safe driver, thus these statistics do not apply to me"). However, whether reasoning is required to discount bad news has not been tested directly. Here, we restrict cognitive resources using a cognitive load (Experiment 1) and a time restriction manipulation (Experiment 3) and find that while these manipulations diminish learning in general, they do not diminish the bias. Furthermore, we also show that the relative neglect of bad news happens the moment new evidence is presented, not when participants are subsequently prompted to state their belief (Experiment 2). Our findings suggest that reasoning is not required for bad news to be discounted as compared to good news.

Introduction

The perception of societal and personal risk is at the heart of public policy, influencing the allocation of resources by governments and organisations (Sunstein, 2005a). While efficient policies likely depend on accurate risk estimates, people's estimates regarding risk can be misaligned with the available evidence (Weinstein, 1980; Taylor & Brown, 1988; Carver *et al.*, 2010; Shepperd *et al.*, 2015; Sharot & Garrett, 2016). For example, providing

* Correspondence to: Department of Psychology, City, University of London, London, UK. Email: kappes.andreas@gmail.com

people with statistics about the risk of smoking, unhealthy eating or drunk driving does not lead to adequate belief adjustments (Weinstein & Klein, 1995; Carver *et al.*, 2010). Most explanations for why people fail to update beliefs in the face of evidence point to motivated reasoning: cognitive processes that attempt to confirm a preferred outcome, belief or judgement (Kunda, 1990; Ditto & Lopez, 1992; Nickerson, 1998; Ditto *et al.*, 2009). Surprisingly, however, while *reasoning* is a central aspect of such models, we are unaware of a direct test on whether it is required for motivated beliefs to emerge.

An alternative mechanism may involve automatic rejection of unwanted information without reasoning (Taber & Lodge, 2006; Mercier & Sperber, 2011). People then will engage only in *post-hoc* rationalisation if prompted (Haidt, 2001; Mercier & Sperber, 2011; Taber & Lodge, 2016). In moral decision-making, for instance, people might have a strong emotional reaction to a certain type of behaviour (e.g., having sex with a dead chicken) that leads them to condemn it even when unable to explain why (Haidt *et al.*, 2000; but see Royzman *et al.*, 2015). Similarly, when confronted with data suggesting one is more likely to become a victim of crime than previously suspected, a negative reaction may be triggered, leading to automatic rejection of the information without reasoning. Here, we test if reasoning is needed to maintain skewed beliefs, focusing on belief updating about personal risks.

Ample research has shown that people are more likely to incorporate good news (i.e., learning the risk of alcohol is lower than they thought) than bad news (i.e., learning it is higher) into prior beliefs (Eil & Rao, 2011; Sharot *et al.*, 2011). Such motivated belief updating has been demonstrated for a range of beliefs relating to personal and societal risks (see Sharot & Garrett, 2016, for an overview), including beliefs about medical risks, safety risks, personal traits, climate change and more (Eil & Rao, 2011; Sharot *et al.*, 2011; Sharot *et al.*, 2012a; Sharot *et al.*, 2012b; Moutsiana *et al.*, 2013; Garrett *et al.*, 2014; Garrett & Sharot, 2014; Korn *et al.*, 2014; Kuzmanovic *et al.*, 2015; Lefebvre *et al.*, 2017). For instance, a study conducted a few months before the 2016 US presidential campaign showed that voters were more inclined to revise their predictions of who was likely to win the elections after receiving good news – that their candidate is leading the polls – than bad news – that the opposing candidate was leading the polls (Tappin *et al.*, 2017). Because different people have different motivations, as in the example above, a relative failure to update beliefs in response to bad news as compared to good news can lead to an increase in belief polarisation within a population (Sunstein *et al.*, 2016).

A reasoning-centric account of motivated belief updating would suggest that people construct reasons for discounting bad news (e.g., a smoker who reads

an article that ties smoking to lung cancer may tell himself: "My grandfather smoked all his life and lived to 100. Thus, I have genes that protect me from the ill effects of smoking") (Lovallo & Kahneman, 2003). Here, reasoning is needed to reject incoming information (Gilbert *et al.*, 1993) and if deliberation is restricted people will update their beliefs equally for good and bad news. However, if bad news can be discounted automatically without deliberation, we should observe motivated belief updating even when cognitive resources are restricted. To test whether deliberation is required for motivated belief updating, we utilised two common manipulations to limit the cognitive resources needed for deliberation: cognitive load and time restriction (Moors & De Houwer, 2006).

Experiment 1

Method

Participants and design

Power calculation. Previous research found large effect sizes for motivated belief updating (Sharot *et al.*, 2011; Sharot *et al.*, 2012a; Sharot *et al.*, 2012b; Moutsiana *et al.*, 2013; Garrett *et al.*, 2014; Garrett & Sharot, 2014; Korn *et al.*, 2014; Kuzmanovic *et al.*, 2015). We assumed a potentially small to medium difference between the load and no-load manipulation ($\eta^2 partial \sim 0.2$). Using G*Power (Faul *et al.*, 2009), we estimated that for an acceptable power of 0.8, we would need about 50 participants, and we recruited 49 participants (22 male, 27 female, M_{age} = 23.56, SD = 4.08).

The study was approved by the University College London Research Ethics Committee. Participants received £10 for participation. We intended to exclude participants with a depression score on the Beck Depression Inventory-II (BDI-II; Beck *et al.*, 1996) of 12 or higher as done previously (Moutsiana *et al.*, 2013). Depression has been shown to eliminate the update bias (Garrett *et al.*, 2014; Korn *et al.*, 2014). However, none of the participants had a BDI-II score that high. The study had a within-subject design with valence of news (good versus bad) and cognitive load (no load versus load) as within-subject factors. Order of condition was randomly assigned. Order did not affect the results, and hence we will not discuss it further.
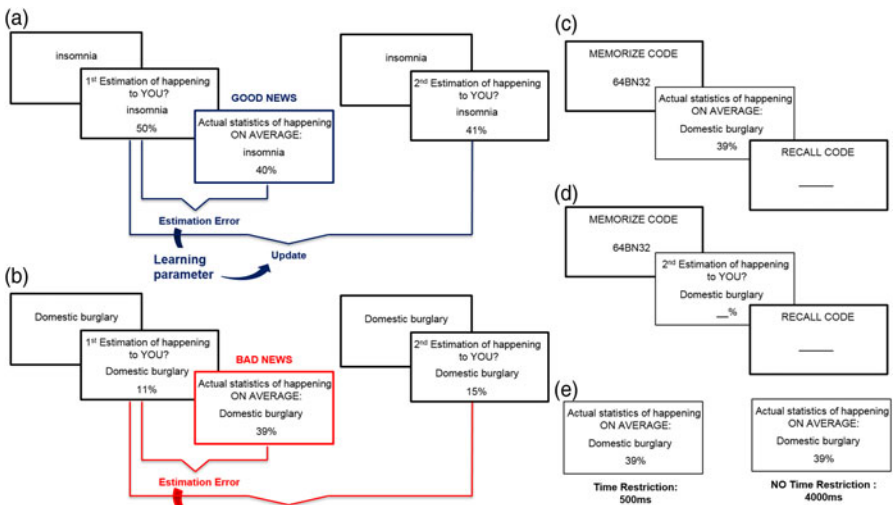
Belief updating task

Stimuli. Eighty short descriptions of negative life events (e.g., passenger in a car accident, home burglary) were presented in random order. The stimuli were ones used in previous research (Sharot *et al.*, 2011; Moutsiana *et al.*,

2013; Garrett & Sharot, 2014). All events were shown to all participants. For each event, the average probability (base rate) of that event occurring at least once to a person living in the same sociocultural environment as the participant was determined from online resources (e.g., Office for National Statistics, Eurostat and PubMed). Our participants were all living in a similar sociocultural environment, so probabilities were the same for all. Very rare or very common events were not included; all event probabilities were between 10% and 70%. To ensure that the range of possible overestimation was equal to the range of possible underestimation, participants were told that the range of probabilities was between 3% and 77%.

Procedure. We used the frequently used 'belief update' procedure (Sharot *et al.*, 2011; Sharot *et al.*, 2012a; Sharot *et al.*, 2012b; Chowdhury *et al.*, 2014; Korn *et al.*, 2014; Kuzmanovic *et al.*, 2015). On each trial, one of the 80 adverse life events was presented in random order for 4 s, and participants were asked to estimate how likely the event was to happen to them in the future. Participants had up to 8 s to respond. They were then presented with the base rate of the event in a demographically similar population for 4 s (see Figure 1). Between each trial, a fixation cross appeared (1 s).

Cognitive load manipulation. To manipulate cognitive resources, participants performed a second task while processing the base rate. Participants were informed that they would complete the task with two different kinds of distraction, since we were interested in studying how people learn under distraction. In the cognitive load condition, participants memorised a password-like digit string (e.g., fA72B6) before the presentation of each base rate and recalled it immediately thereafter (e.g., see DeShon *et al.*, 1996; Conway & Gawronski, 2013, for similar procedures). In order to ensure that participants were motivated to do so, we paid them an additional £2 if they correctly recalled at least 60% of passwords. All but one participant did so. On average, participants correctly remembered 88% (SD = 8%) of passwords. In the no load condition, participants saw the same digit strings as in the load condition, but were asked to ignore them.

In a second session immediately after the first, participants were asked again to provide estimates of their likelihood of encountering the same events, so that we could assess how they updated their estimates in response to the information presented. After participants finished the task, participants rated all stimuli on prior experience (for the question "Has this event happened to you before?" the responses ranged from 1 [never] to 6 [very often]), familiarity (for the question "Regardless if this event has happened to you before, how familiar do you feel it is to you from TV, friends, movies and so on?" the

Figure 1. Belief updating task. (a) Example of a good news trial – the first estimate is higher than the average likelihood displayed. The estimation error is then calculated by subtracting the first estimate from the average likelihood and the update is calculated by subtracting the first estimate from the second estimate. The learning parameter indicates how well the estimation error predicted subsequent update. (b) Example of a bad news trial – the first estimate is lower than the average likelihood. The estimation error is then calculated by subtracting the average likelihood from the first estimate and the update is calculated by subtracting the second estimate from the first estimate. (c) Cognitive load in Experiment 1 was induced by asking participants to memorise a code before observing the average information and recalling it immediately thereafter. (d) Cognitive load in Experiment 2 was induced by asking participants to memorise a code before entering their second estimate and recalling it immediately thereafter. (e) In Experiment 3, information was presented for either 500 ms (time restriction) or for 4000 ms (no time restriction).

responses ranged from 1 [not at all familiar] to 6 [very familiar]) and negativity (for the question "How negative would this event be for you?" the responses ranged from 1 [not negative at all] to 6 [very negative]). To test memory for the information presented, subjects were asked at the end to provide the actual probability previously presented of each event. Data of the control variables from one participant were lost due to a computer crash.

Data analysis. Data analysis was equivalent to previous research (Sharot *et al.*, 2011, 2012a, 2012b; Moutsiana *et al.*, 2013; Chowdhury *et al.*, 2014; Garrett *et al.*, 2014). For each event, an estimation error term was calculated as the difference between participants' first estimate and the corresponding base rate: estimation error = first estimation – base rate presented. Estimation errors were positive for overestimations and negative for underestimations. When participants initially overestimated the probability of the adverse event relative to the average probability, they received good news (i.e., the negative event is less likely to happen than estimated; Figure 1(a)). By contrast, when participants underestimated the probability of the event relative to the average probability, they received bad news (i.e., the negative event is more likely to happen than estimated; Figure 1(b)). Therefore, for each participant, trials were classified according to whether the participant initially overestimated or underestimated the probability of the event (i.e., according to whether estimation errors were positive or negative).

For each trial and each participant, we then estimated update as follows: good news update = first estimation – second estimation; bad news update = second estimation – first estimation. Thus, positive updates indicate a change towards the base rate and negative updates indicate a change away from the base rate. For each participant, we then averaged updates scores across trials for which good news was presented and separately for which bad news was presented. To test the strength of association between estimation errors and updates (i.e., learning parameter), Pearson correlation coefficients were calculated separately for good and bad news trials within each participant.

Results and discussion

Dual-task paradigms test the automaticity (efficiency or independence of deliberation) of the main task (here, the belief updating task) by letting participants simultaneously perform a second task (here, remembering passwords) (Moors & De Houwer, 2006). If motivated belief updating is automatically achieved, then we should see no difference in the update bias between the cognitive load condition and the no load condition. However, if deliberation is needed in order to bias the updating, then we should see an update bias only in the no load condition, but not in the cognitive load condition.

In order to test if motivated belief updating is automatic or deliberate, we entered update as the dependent variable into a repeated measure of analysis, and as independent variables the valence of the information (good versus bad news) and load manipulation (cognitive load versus no load) (Figure 2). We controlled for differences in estimation errors for good versus bad news in each condition to ensure that differences in learning for good versus bad
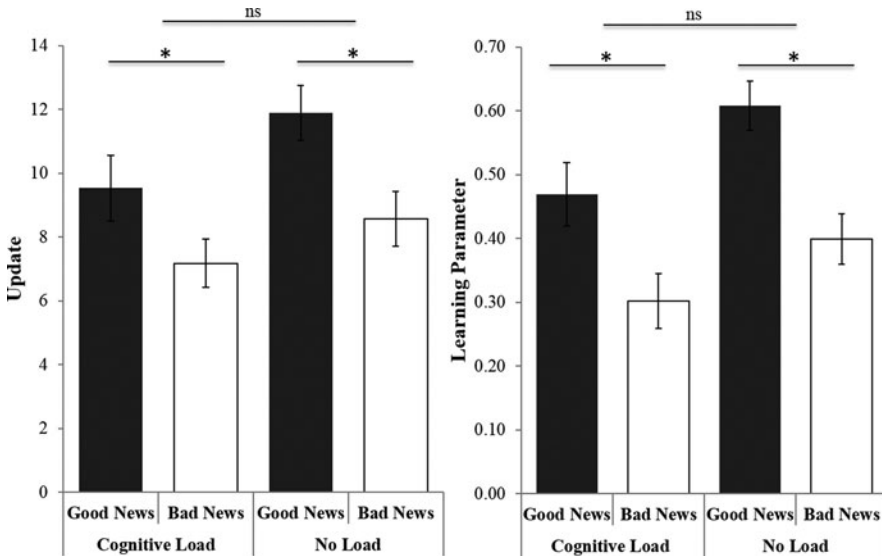
Figure 2. Cognitive load did not affect valence-dependent asymmetric updating (left) nor asymmetric learning parameters (right). Participants updated their belief to a greater extent after receiving good news compared to bad news and had higher learning parameters. Error bars represent standard errors of the mean. *$p < 0.05$. ns = non-significant.

news do not reflect differences in the initial estimates (i.e., prior beliefs) between conditions or mean estimation errors (Garrett & Sharot, 2017). Note that we find the same results when not controlling for differences in estimation errors. First, we found a main effect for the load manipulation, $F(1,46) = 6.59$, $p = 0.01$, $\eta^2 partial = 0.09$. Participants updated less in the cognitive load condition (M = 8.35, SE = 0.62) compared to the no load condition (M = 10.23, SE = 0.58). The main effect indicates that our cognitive load manipulation was successful, reducing belief updating when a strain was put on the available cognitive resources.

As predicted, we found a main effect for valence, $F(1,46) = 9.61$, $p = 0.003$, $\eta^2 partial = 0.17$. Participants updated their beliefs more in response to good news (M = 10.71, SE = 0.76) than bad news (M = 7.87, SE = 0.66). Importantly, we did not find an interaction effect between condition and valence, $F(1,46) = 0.17$, $p = 0.68$, $\eta^2 partial = 0.004$. The effect size of $\eta^2 partial = 0.004$ for the interaction effect is well below a small effect ($\sim\eta^2 partial = 0.02$).

We repeated the same analysis using the learning parameter as the dependent variable. Again, we found a main effect for load, $F(1,46) = 8.88$, $p = 0.005$, $\eta^2 partial = 0.16$, showing that the learning parameter under load was

significantly smaller (M = 0.39, SE = 0.037) than under no load (M = 0.50, SE = 0.033). Our manipulation effectively interfered with the relationship between the estimation error and the subsequent update. We additionally found the predicted effect of valence (good versus bad news), $F(1,46) = 20.24$, $p < 0.0001$, $\eta^2 partial = 0.31$; participants' learning parameters were significantly stronger in response to good news (M = 0.54, SE = 0.35) than bad news (M = 0.36, SE = 0.03). Finally, as for update, we did not find a significant interaction effect of condition and valence on the learning parameter, $F(1,46) = 0.08$, $p = 0.77$, $\eta^2 partial = 0.002$. Again, the effect size was well below a small effect size, suggesting that such bias does not require deliberation.

Next, we tested whether the control variables (memory, vividness, experience, negativity and arousal) could explain either the main effect of valence and/or the main effect of load we observed. Only for participants' memory did we find a main effect of valence, $F(1,47) = 7.62$, $p = 0.008$, $\eta^2 partial = 0.13$ and load $F(1,47) = 6.37$, $p = 0.015$, $\eta^2 partial = 0.14$. We then repeated the main analyses reported above, this time controlling for the differences in memory between good and bad news, as well as load versus no load. For belief update, we found that when controlling for difference in memory, the main effect of valence remained significant, $F(1,45) = 5.81$, $p = 0.02$, $\eta^2 partial = 0.11$, while the main effect of load disappeared, $F(1,45) = 0.82$, $p = 0.37$, $\eta^2 partial = 0.01$. For learning parameters, we again found that when controlling for difference in memory, the main effect of valence remained significant, $F(1,45) = 8.81$, $p = 0.005$, $\eta^2 partial = 0.16$, while the main effect of load disappeared, $F(1,45) = 0.95$, $p = 0.335$, $\eta^2 partial = 0.021$. These findings suggests that the cognitive load manipulation interfered with participants' ability to successful encode the average information, but not with their bias.

One limitation of the presented analysis is that classic frequentist analyses (testing the significance of differences, for instance) does not provide evidence for a null effect, as we reported here. So, we complemented our results with Bayesian analyses, which can provide evidence in favour of the null hypothesis. In particular, we performed Bayesian t-tests for accepting or rejecting the null hypothesis, comparing the bias in updating and learning parameter under load versus no load to generate Bayes factors (BF_{10}) (Rouder *et al.*, 2009). We used JASP (Love *et al.*, 2015). A BF_{10} greater than 1 provides evidence for rejecting the null hypothesis, a BF_{10} smaller than 1 provides evidence for the null hypothesis. When comparing the difference between update from good versus bad news when participants learned under cognitive load versus no cognitive load, we find $BF_{10} = 0.192$, suggesting substantial evidence for the null hypothesis (Rouder *et al.*, 2009). Repeating the same analysis for the learning

parameter produced a $BF_{10} = 0.179$, again providing substantial evidence for the null hypothesis.

In summary, we found evidence suggesting that the difference in updating from good and bad news does not require deliberation. In particular, while our load manipulation successfully affected learning – participants learned less under load than under no load as indicated by lower updates and learning parameters – it did not affect the bias towards good news. Participants in both conditions showed the same difference between learning from good versus bad news.

Experiment 2

Experiment 1 suggests that restricting cognitive resources at the time evidence is presented does not eliminate motivated belief updating. In Experiment 2, we ask whether cognitive load at the time participants entered their updated beliefs will affect their bias (see Figure 1(d)). We assumed belief updating occurs when evidence is presented, not when new estimates are elicited, and thus predicted no effect of the manipulation on belief updating. However, it is plausible that a process of rationalisation, which contributes to the update bias, occurs when people are asked to provide their new estimates. If this is the case, we should observe no bias and a reduced bias in belief updating under cognitive load.

Method

Participants and design

Power calculation. We assumed a potentially small difference between load and no-load manipulation ($\eta^2 partial \sim 0.1$). Using G*Power (Faul *et al.*, 2009), we estimated that for a power of 0.9, we would need about 60 participants, and we recruited 68 participants (23 male, 45, female, $M_{age} = 23.02$, $SD = 4.51$). The study was approved by the University College London Research Ethics Committee. Participants received £10 as a participation fee. We exclude 12 participants with a depression score on the BDI-II (Beck *et al.*, 1996) of 12 or higher, indicating moderate depression, as done previously (Moutsiana *et al.*, 2013). The study had a within-subject design with valence of news (good versus bad) and cognitive load (no load versus load) as within-subject factors. Order of condition was random and did not affect the results.

Cognitive load manipulation. The belief updating task was the same one used in Experiment 1. However, this time we manipulated the availability of cognitive resources when participants were prompted to enter their updated

beliefs (see Figure 1(d)). In the cognitive load condition, participants memorised a password-like digit string (e.g., fA72B6) before they re-entered their belief about the likelihood of the negative life event happening to them and recalled it immediately after.

Results and discussion

We entered update as the dependent variable into a repeated-measures analysis of variance (ANOVA), and as independent variables the valence of the information (good versus bad news) and manipulation (cognitive load versus no cognitive load) (Figure 3). Again, we controlled for differences in estimation errors for good versus bad news in each condition to ensure that differences in learning for good versus bad news did not reflect differences in the initial estimates (i.e., prior beliefs) between conditions or mean estimation errors (Garrett & Sharot, 2017). We found a main effect for valence, $F(1,52) = 8.24$, $p = 0.006$, $\eta^2_{partial} = 0.14$. As expected, participants updated their beliefs more in response to good news (M = 11.49, SE = 0.64) than bad news (M = 9.02, SE = 0.62). As predicted, we did not find a main effect for the load manipulation, $F(1,52) = 1.19$, $p = 0.28$. Participants on average updated as much under
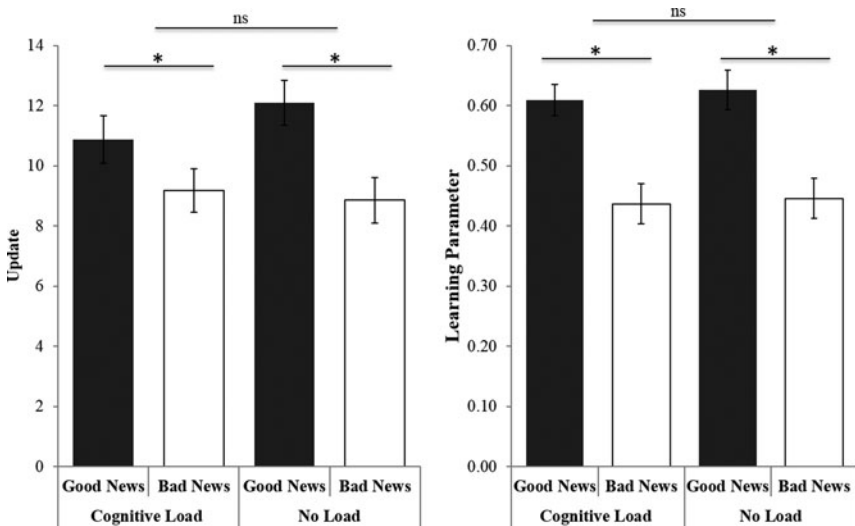
Figure 3. Cognitive load manipulation did not affect updating (left) nor learning parameters (right) after receiving either good or bad news. Participants updated their belief to a greater extent following good news compared to bad news and had higher learning parameters. Error bars represent standard errors of the mean. *$p < 0.05$. ns = non-significant.

cognitive load (M = 10.03, SE = 0.49) as under no load (M = 10.47, SE = 0.55). Given that the same cognitive load manipulation influenced learning when applied during the presentation of the new information, the lack of main effect here suggests that learning already took place before participants entered their second estimate. Finally, we did not find an interaction effect between condition and valence, $F(1,52) = 0.70$, $p = 0.40$, $\eta^2 partial = 0.01$. When repeating our analysis with learning parameter as the dependent variable, we found the same results. In particular, we found the expected main effect for valence, $F(1,52) = 32.46$, $p < 0.001$, $\eta^2 partial = 0.38$, no main effect for load manipulation, $F(1,52) = 0.56$, $p = 0.46$, and no interaction effect, $F(1,52) = 0.003$, $p = 0.96$.

Taken together, the findings of Experiments 1 and 2 suggest that learning took place when participants received new information. Furthermore, Experiment 1 suggests that the update bias does not require deliberation. In order to conceptually replicate our finding, we next used a time manipulation to restrain cognitive resources. Specifically, participants had either 0.5 s to process the information (i.e., time restriction) or 4 s (no time restriction).

Experiment 3: Asymmetric belief updating under time restriction

Method

Participants and design

Power calculation. We assumed a potentially small to medium difference between time restriction and no time restriction manipulation ($\eta^2 partial \sim 0.2$). Using G*Power (Faul *et al.*, 2009), we estimated that for an acceptable power of 0.8, we would need about 50 participants, and we recruited 55 participants (26 male, 29 female, $M_{age} = 22.12$, SD = 3.91). Participants received £10 as a participation fee. The study was approved by the University College London Research Ethics Committee. Three participants with a depression score of 12 or higher on the BDI-II (Beck *et al.*, 1996) indicating moderate depression were excluded. The study had a 2×2 within-subjects design with factors valence of news (good versus bad news) and time restriction (no time restriction versus time restriction). Condition order was random and did not affect the results.

Time restriction manipulation. The belief updating task had the same structure as in Experiment 1. However, this time, we manipulated the time participants had to process the evidence (e.g., see Suter & Hertwig, 2011; Rand *et al.*, 2012, for similar manipulations). In particular, in the time restriction condition, participants were presented with the information for only 0.5 s. Such a

limited amount of time is not sufficient for conscious deliberation (Suter & Hertwig, 2011). In contrast, in the no time restriction condition, participants were presented with the information for 4 s. This is an equal duration to that in Experiment 1 and allows conscious deliberation.

Results and discussion

Time restriction manipulations are based on a similar assumption as dual-task manipulations. Under time restriction, participants do not have enough cognitive resources available to deliberate and hence, if asymmetric belief updating is dependent on such deliberation, it should disappear. However, if asymmetric belief updating is automatic, we should find no difference between the time restriction and the no time restriction conditions.

We entered update as the dependent variable into a repeated-measures ANOVA, and as independent variables the valence of the information (good versus bad news) and time manipulation (time restriction versus no time restriction) (Figure 4). Again, we controlled for differences in estimation errors for good versus bad news in each condition to ensure that differences in learning for good versus bad news did not reflect differences in the initial estimates (i.e., prior beliefs) between conditions or mean estimation errors (Garrett & Sharot, 2017). First, we found a main effect for time manipulation, $F(1,49) = 4.76$, $p = 0.034$, $\eta^2 partial = 0.09$. Participants on average updated less in the time restriction condition (M = 9.7, SE = 0.62) compared to the no time restriction condition (M = 10.98, SE = 0.53). The main effect indicates that our time restriction manipulation was successful, reducing learning when time reduced available cognitive resources. We also found a main effect for valence, $F(1,49) = 4.36$, $p = 0.042$, $\eta^2 partial = 0.14$. As expected, participants updated their beliefs more in response to good news (M = 12.01, SE = 0.79) than bad news (M = 8.67, SE = 1.06). Importantly, we again did not find an interaction effect between condition and valence, $F(1,49) = 0.04$, $p = 0.83$, $\eta^2 partial = 0.001$. The effect size is well below a small effect, indicating a lack of effect.

We repeated the same analysis, this time using the learning parameter as the dependent variable. Again, we found the main effect of valence, $F(1,49) = 25.52$, $p < 0.0001$, $\eta^2 partial = 0.32$, showing that the learning parameter related to positive news was significantly larger (M = 0.56, SE = 0.029) than the one related to negative news (M = 0.39, SE = 0.029). However, we did not find a main effect for the time manipulation, $F(1,49) = 0.55$, $p = 0.46$, $\eta^2 partial = 0.01$. Here, there was an interaction effect between valence and time manipulation, $F(1,49) = 4.24$, $\eta^2 partial = 0.08$. However, in contradiction to a deliberative account of the update bias, the difference between the learning
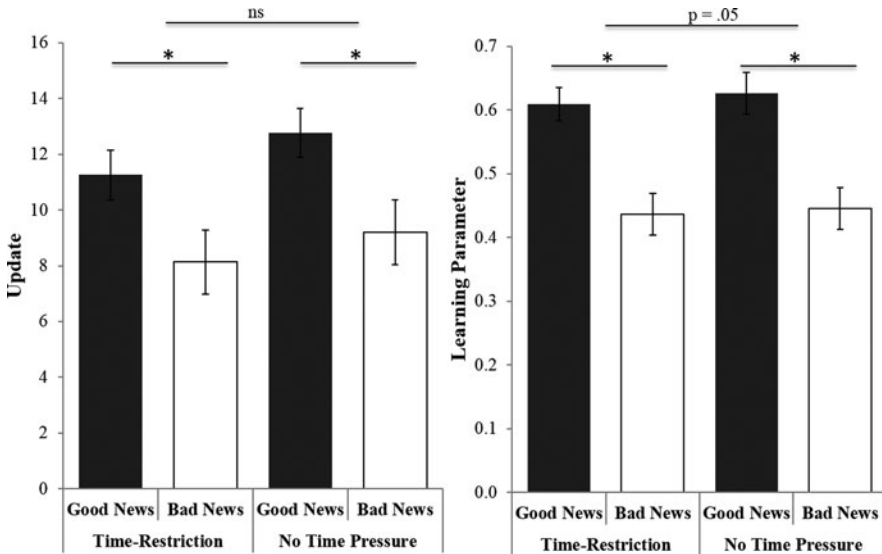
Figure 4. The time restriction manipulation did not affect the difference between updating (left) in response to good news versus bad news nor the difference in learning parameters (right). Participants updated their beliefs to a greater extent in response to good news versus bad news and had higher learning parameters. Error bars represent standard errors of the mean. *$p < 0.05$. ns = non-significant.

parameters from good and bad news was more pronounced under time restriction ($M_{diff} = 0.22$) than under no time restriction ($M_{diff} = 0.11$). However, given that we did not find the same interaction effect between valence and condition for update, and nor did we find such an interaction effect in Experiment 1, this result must be interpreted cautiously.

We then used Bayesian statistics to find support in favour of the hypothesis of an automatic learning bias. When comparing the difference between updating from good versus bad news under time restriction with the difference under no time restriction, we found a $BF_{10} = 0.16$, suggesting substantial evidence for the null hypothesis.

General discussion

We set out to test whether motivated belief updating is dependent on deliberation. Using two methods for limiting controlled processing – cognitive load and time manipulation – we found evidence that deliberation is not required for a bias to emerge by which people update their belief more in response to

good news than bad news. In particular, while our manipulations reduced the amount by which beliefs were updated, suggesting that cognitive resources were successfully restricted, they did not influence the bias. This suggests that discounting bad news can occur automatically when cognitive resources are limited and ability to reason is restricted. Our findings also suggest that belief updating takes place when evidence is presented rather than at a later stage. In particular, belief updating was impaired when cognitive resources were restricted at the time evidence was presented, but not at the time the new belief was elicited.

Our results support the suggestion that motivated cognition might be driven by automatic rather than deliberate processes (e.g., Sunstein, 2005b; Weber, 2006; Taber & Lodge, 2016). However, it is important to note that we examined only one paradigm, looking specifically at beliefs regarding negative life events. It is possible that deliberation does play a role in other situations (Hamilton *et al.*, 2012; Kahan *et al.*, 2017). Our study highlights that in order to claim as much, cognitive resources need to be manipulated.

An automatic failure to integrate bad news relative to good news into one's beliefs could be an innate tendency, potentially reflecting an evolutionary advantage (Haselton & Nettle, 2006; McKay & Dennett, 2009; Johnson & Fowler, 2011). Maintaining positively skewed beliefs about oneself and one's future has utility for the individual, reducing stress and contributing positively to physical and mental health (Taylor & Brown, 1988; Taylor *et al.*, 2000; Hernandez *et al.*, 2015). In addition, having an overly positive self-view may help convince others of one's value (Smith *et al.*, 2017). Similarly, evolutionary computational models suggest that being overly optimistic about one's own capabilities helps to outdo the competition in environments where the rewards of winning resources (e.g., food or mating partners) outweigh the potential costs of competing (Johnson & Fowler, 2011).

Yet, even if motivated updating of beliefs provides an advantage to the individual, automatic processes do not have to be innate, but rather the result of a learning process (Dolan & Dayan, 2013). With respect to neglecting bad news, children might have been rewarded for thinking positively about themselves and their future from an early age and discouraged from a pessimistic, self-critical worldview (Carver & Scheier, 2014). Parental instruction and praise might have established an automatic habit of rejecting bad news. A learned rejection of bad news would help explain why people in cultures that reward a self-critical focus lack positively skewed self-beliefs (Heine *et al.*, 1999). More research is needed to test whether the automatic tendency to dismiss bad news is innate, overlearned or a combination of both.

References

Beck, A. T., R. A. Steer, R. Ball, and W. Ranieri (1996), 'Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients', *Journal of Personality Assessment*, **67**(3): 588–597.

Carver, C. S., and M. F. Scheier (2014), 'Dispositional Optimism', *Trends in Cognitive Sciences*, **18**(6): 293–299.

Carver, C. S., M. F. Scheier, and S. C. Segerstrom (2010), 'Optimism', *Clinical Psychology Review*, **30**(7): 879–889.

Chowdhury, R., T. Sharot, T. Wolfe, E. Düzel, and R. J. Dolan (2014), 'Optimistic update bias increases in older age', *Psychological Medicine*, **44**(09): 2003–2012.

Conway, P., and B. Gawronski (2013), 'Deontological and utilitarian inclinations in moral decision making: a process dissociation approach', *Journal of Personality and Social Psychology*, **104**(2): 216–235.

DeShon, R. P., K. G. Brown, and J. L. Greenis (1996), 'Does self-regulation require cognitive resources? Evaluation of resource allocation models of goal setting', *The Journal of Applied Psychology*, **81**(5): 595–608.

Ditto, P. H., and D. F. Lopez (1992), 'Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions', *Journal of Personality and Social Psychology*, **63**(4): 568.

Ditto, P. H., D. A. Pizarro, and D. Tannenbaum (2009), Chapter 10 Motivated Moral Reasoning. In B. H. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 50, pp. 307–338), Academic Press. Retrieved from http://www.sciencedirect.com/science/article/pii/S0079742108004106

Dolan, R. J., and P. Dayan (2013), 'Goals and Habits in the Brain', *Neuron*, **80**(2): 312–325.

Eil, D., and J. M. Rao (2011), 'The Good News-Bad News Effect: Asymmetric Processing of Objective Information about Yourself', *American Economic Journal: Microeconomics*, **3**(2): 114–138.

Faul, F., E. Erdfelder, A. Buchner, and A.-G. Lang (2009), 'Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses', *Behavior Research Methods*, **41**(4): 1149–1160.

Garrett, N., and T. Sharot (2014), 'How robust is the optimistic update bias for estimating self-risk and population base rates'? Retrieved from http://dx.plos.org/10.1371/journal.pone.0098848

Garrett, N., and T. Sharot (2017), 'Optimistic update bias holds firm: Three tests of robustness following Shah *et al.*', *Consciousness and Cognition*, **50**(Supplement C): 12–22.

Garrett, N., T. Sharot, P. Faulkner, C. W. Korn, J. P. Roiser, and R. J. Dolan (2014), 'Losing the rose tinted glasses: neural substrates of unbiased belief updating in depression', *Frontiers in Human Neuroscience*, **8**. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4147849/

Gilbert, D. T., R. W. Tafarodi, and P. S. Malone (1993), 'You can't not believe everything you read', *Journal of Personality and Social Psychology*, **65**(2): 221–233.

Haidt, J. (2001), 'The emotional dog and its rational tail: A social intuitionist approach to moral judgment', *Psychological Review*, **108**(4): 814–834.

Haidt, J., F. Bjorklund, and S. Murphy (2000), 'Moral dumbfounding: When intuition finds no reason', *Unpublished Manuscript, University of Virginia*. Retrieved from http://www.faculty.virginia.edu/haidtlab/articles/manuscripts/haidt.bjorklund.working-paper.when%20intuition%20finds%20no%20reason.pub603.doc

Hamilton, L. C., M. J. Cutler, and A. Schaefer (2012), 'Public knowledge and concern about polar-region warming', *Polar Geography*, **35**(2): 155–168.

Haselton, M. G., and D. Nettle (2006), 'The Paranoid Optimist: An Integrative Evolutionary Model of Cognitive Biases', *Personality and Social Psychology Review*, **10**(1): 47–66.

Heine, S. J., D. R. Lehman, H. R. Markus, and S. Kitayama (1999), 'Is there a universal need for positive self-regard'?, *Psychological Review*, **106**(4): 766.

Hernandez, R., K. N. Kershaw, J. Siddique, J. K. Boehm, L. D. Kubzansky, A. Diez-Roux, and D. M. Lloyd-Jones (2015), 'Optimism and cardiovascular health: multi-ethnic study of atherosclerosis (MESA)', *Health Behavior and Policy Review*, **2**(1): 62–73.

Johnson, D. D. P., and J. H. Fowler (2011), 'The evolution of overconfidence', *Nature*, **477**(7364): 317–320.

Kahan, D. M., E. Peters, E. C. Dawson, and P. Slovic (2017), 'Motivated numeracy and enlightened self-government'. *Behavioural Public Policy*, **1**(1): 54–86.

Korn, C. W., T. Sharot, H. Walter, H. R. Heekeren, and R. J. Dolan (2014), 'Depression is related to an absence of optimistically biased belief updating about future life events', *Psychological Medicine*, **44**(3): 579–592.

Kunda, Z. (1990), 'The case for motivated reasoning', *Psychological Bulletin*, **108**(3): 480.

Kuzmanovic, B., A. Jefferson, and K. Vogeley (2015), 'Self-specific Optimism Bias in Belief Updating Is Associated with High Trait Optimism', *Journal of Behavioral Decision Making*, **28**(3): 281–293.

Lefebvre, G., M. Lebreton, F. Meyniel, S. Bourgeois-Gironde, and S. Palminteri (2017), 'Behavioural and neural characterization of optimistic reinforcement learning', *Nature Human Behaviour*, **1**: 0067.

Lovallo, D., and D. Kahneman (2003), 'Delusions of success', *Harvard Business Review*, **81**(7): 56–63.

Love, J., R. Selker, M. Marsman, T. Jamil, A. J. Verhagen, A. Ly, and others. (2015), 'JASP (Version 0.7)', *Computer Software*.

McKay, R. T., and D. C. Dennett (2009), 'The evolution of misbelief', *Behavioral and Brain Sciences*, **32**(06): 493–510.

Mercier, H., and D. Sperber (2011), 'Why do humans reason? Arguments for an argumentative theory', *Behavioral and Brain Sciences*, **34**(02): 57–74.

Moors, A., and J. De Houwer (2006), 'Automaticity: a theoretical and conceptual analysis', *Psychological Bulletin*, **132**(2): 297.

Moutsiana, C., N. Garrett, R. C. Clarke, R. B. Lotto, S.-J. Blakemore, and T. Sharot (2013), 'Human development of the ability to learn from bad news', *Proceedings of the National Academy of Sciences*, **110**(41): 16396–16401.

Nickerson, R. S. (1998), 'Confirmation bias: A ubiquitous phenomenon in many guises', *Review of General Psychology*, **2**(2): 175–220.

Rand, D. G., J. D. Greene, and M. A. Nowak (2012), 'Spontaneous giving and calculated greed', *Nature*, **489**(7416): 427–430.

Rouder, J. N., , P. L. Speckman, D. Sun, R. D. Morey, and G. Iverson (2009), 'Bayesian t tests for accepting and rejecting the null hypothesis', *Psychonomic Bulletin & Review*, **16**(2): 225–237.

Royzman, E. B., K. Kim, and R. F. Leeman (2015), 'The curious tale of Julie and Mark: Unraveling the moral dumbfounding effect', *Judgment and Decision Making*, **10**(4): 296.

Sharot, T., and N. Garrett (2016), 'Forming Beliefs: Why Valence Matters', *Trends in Cognitive Sciences*, **20**(1): 25–33.

Sharot, T., M. Guitart-Masip, C. W. Korn, R. Chowdhury, & R. J. Dolan (2012a), 'How dopamine enhances an optimism bias in humans', *Current Biology*, **22**(16): 1477–1481.

Sharot, T., R. Kanai, D. Marston, C. W. Korn, G. Rees, and R. J. Dolan (2012b), 'Selectively altering belief formation in the human brain', *Proceedings of the National Academy of Sciences*, **109**(42): 17058–17062.

Sharot, T., C. W. Korn, and R. J. Dolan (2011), 'How unrealistic optimism is maintained in the face of reality', *Nature Neuroscience*, **14**(11): 1475–1479.

Shepperd, J. A., E. A. Waters, N. D. Weinstein, and W. M. P. Klein (2015), 'A Primer on Unrealistic Optimism', *Current Directions in Psychological Science*, **24**(3): 232–237.

Smith, M. K., R. Trivers, and W. von Hippel (2017), 'Self-deception facilitates interpersonal persuasion'. *Journal of Economic Psychology*, **63**: 93–101.

Sunstein, C. R. (2005a), *Laws of fear: Beyond the precautionary principle* (Vol. 6). Cambridge University Press. Retrieved from https://books.google.co.uk/books?hl=en&lr=&id=-OVUQPirb6cC&oi=fnd&pg=PP1&dq=Laws+of+Fear:+Beyond+the+Precautionary+Principle&ots=7F3BKUco9h&sig=L4RE6izNxStxQUjKCXh8fetK2JE

Sunstein, C. R. (2005b), 'Misfearing: A reply', *Harvard Law Review*, **119**: 1110.

Sunstein, C. R., S. Bobadilla-Suarez, S. C. Lazzaro, and T. Sharot (2016), *How People Update Beliefs about Climate Change: Good News and Bad News* (SSRN Scholarly Paper No. ID 2821919). Rochester, NY: Social Science Research Network. Retrieved from https://papers.ssrn.com/abstract=2821919.

Suter, R. S., and R. Hertwig (2011), 'Time and moral judgment', *Cognition*, **119**(3): 454–458.

Taber, C. S., and M. Lodge (2006), 'Motivated Skepticism in the Evaluation of Political Beliefs', *American Journal of Political Science*, **50**(3): 755–769.

Taber, C. S., and M. Lodge (2016), 'The illusion of choice in democratic politics: the unconscious impact of motivated political reasoning', *Political Psychology*, **37**(S1): 61–85.

Tappin, B. M., L. van der Leer, and R. T. McKay (2017), 'The Heart Trumps the Head: Desirability Bias in Political Belief Revision', Retrieved from http://psycnet.apa.org/psycarticles/2017-23363-001

Taylor, S. E., and J. D. Brown (1988), 'Illusion and well-being: A social psychological perspective on mental health', *Psychological Bulletin*, **103**(2): 193–210.

Taylor, S. E., M. E. Kemeny, G. M. Reed, J. E. Bower, and T. L. Gruenewald (2000), 'Psychological resources, positive illusions, and health', *American Psychologist*, **55**(1): 99.

Weber, E. U. (2006), 'Experience-Based and Description-Based Perceptions of Long-Term Risk: Why Global Warming does not Scare us (Yet)', *Climatic Change*, **77**(1–2): 103–120.

Weinstein, N. D. (1980), 'Unrealistic optimism about future life events', *Journal of Personality and Social Psychology*, **39**(5): 806–820.

Weinstein, N. D., and W. M. Klein (1995), 'Resistance of personal risk perceptions to debiasing interventions', *Health Psychology*, **14**(2): 132–140.