Autism, optimism and positive events: Evidence against a general optimistic bias

Adam J. L. Harris (adam.harris@ucl.ac.uk)
Department of Cognitive, Perceptual & Brain Sciences, University College London, Gower Street, London, UK

Punit Shah (ps00113@surrey.ac.uk) and Caroline Catmur (c.catmur@surrey.ac.uk)
Department of Psychology, University of Surrey, Guildford, UK

Geoffrey Bird¹ (geoff.bird@kcl.ac.uk) and Ulrike Hahn (u.hahn@bbk.ac.uk)
Department of Psychological Sciences, Birkbeck College, University of London, Malet Street, London, UK

Abstract
The conclusion that people are optimistic concerning personal risk does not have a sound evidential basis. Following Harris and Hahn’s (2011) critique of unrealistic optimism research, we consider the evidence from a recent series of high profile neuroscience papers. We demonstrate that the methods used are fundamentally flawed. A simulation and an empirical comparison of autism spectrum condition participants with typical adults confirm that we have learnt nothing about optimism from these studies.

Keywords: Optimism; human rationality; belief updating; statistical artifact

Introduction
It is a long-established ‘fact’ that human judgment is biased when estimating personal risk. It has been argued that people underestimate their risk of encountering negative events (compared to their estimate of the average person’s risk), and overestimate their chances of encountering positive events (e.g., Weinstein, 1980). This pattern of optimistic self-estimates has been termed ‘unrealistic optimism’. The majority of evidence for unrealistic optimism originated from research employing what we will term the ‘comparison method’. Participants are asked to directly (Weinstein, 1980) or indirectly (e.g., van der Velde, van der Pligt, & Hooykas, 1994; for a review, see Helweg-Larsen & Shepperd, 2001) compare their chance of experiencing a negative life event with that of the average individual. It is argued that the existence of unrealistic optimism is supported by a pattern of results where the mean of participants’ estimates of self-risk is significantly lower than the mean of their estimates of the average person’s risk. However, a recent analysis of studies using the comparison method casts doubt over the existence of unrealistic optimism. Harris and Hahn (2011) showed that perfectly rational (non-biased) agents generated personal risk estimates that would be classified as unrealistically optimistic on the comparison method for purely statistical reasons. Unfortunately, the comparison method has been used in the vast majority of unrealistic optimism studies to date, meaning that rather than being a characteristic of healthy human thought (e.g., Sharot, 2012; Taylor & Brown, 1988), ‘unrealistic optimism’ may instead be a statistical artifact resulting from a flawed methodology.

Identifying Optimism: A New Development
Unrealistic optimism is widely regarded as a central cognitive bias, yet the fact that almost all past research is based on the comparison method means that there is very little empirical evidence for its existence. The frailties associated with the comparison method have, however, only recently been identified, and common belief amongst both researchers and the general public is that risk estimates are characterised by an optimism bias (e.g., Sharot, 2012).

With this position as their starting point, Sharot and colleagues (Sharot, Guitart-Masip et al., 2012; Sharot, Kanai, et al., 2012; Sharot, Korn, & Dolan, 2011) have begun an investigation into the neural underpinnings of the phenomenon. Sharot et al. (2011) claimed that people maintain an optimistic view of the future through selectively incorporating new desirable information into their beliefs more than new undesirable information. If this result is robust, it seems to shift the body of evidence back in favor of an optimistic bias in human likelihood judgments.

The Update Method We shall refer to the method introduced in Sharot et al. (2011; see also, Sharot, Guitart-Masip, et al., 2012; Sharot, Kanai, et al., 2012) as the update method (Figure 1). Participants estimate their chance of experiencing each of a series of negative events. They are subsequently given the probabilities with which these negative events are experienced by the average individual (base rates). Subsequently, participants re-estimate their chances of experiencing the same negative life events. The critical variable is the degree to which participants update their personal risk estimates in light of the given information about the average person’s risk. Updates are compared across two types of trial: trials where the base rate of the negative event was lower than the participant’s initial estimate (‘desirable information’), and trials where it was higher (‘undesirable information’). Participants are found to update their estimates more in response to desirable

¹ Geoffrey Bird is now at Institute of Psychiatry, King’s College London.
information, thus enabling the preservation of an unrealistically optimistic view of the future.

Figure 1: The update method.

**Overview**

The present paper proceeds in two stages. In the first stage, we present a formal critique of the update method. We subsequently demonstrate (Study 1) that, as a result of the identified flaws, the update method gives rise to a seemingly optimistic pattern of belief updating in a population of rational Bayesian agents. Finally, in anticipation of a counter-argument that the demonstration of neural moderators of the effect (e.g., Sharot, Kanai, et al., 2012) provides further evidence for its existence, we present results from a study comparing autism spectrum condition (ASC) participants with typical adult (TA) controls. With a suitable addition to the update method, we demonstrate that we can infer nothing about the neural correlates of optimism from the results currently present in the literature.

**What’s Wrong With the Update Method?**

There are two distinct ways in which we might receive new information relevant to our individual risk. We may receive new information about the base rate, or we may receive diagnostic information that pertains to us personally (e.g., genetic vulnerability). Both types of information are relevant, both can be desirable or undesirable, and both should be combined via Bayes’ Theorem to provide our best estimate of risk (e.g., Kahneman & Tversky, 1973), as is illustrated in the following example.

**Normative Risk Updating** 55-year-old Tim estimates that the average 55-year-old’s risk of contracting heart disease (HD) (the base rate) is 20%. In the absence of any other information, Tim’s best estimate of his own likelihood of contracting HD is 20%.

If Tim possesses any diagnostic information that differentiates his risk from the average person’s, he should normatively combine the base rate with this diagnostic information. For example, if he does not have a family history of HD, his risk is lower than the average person’s. Bayes’ Theorem prescribes how this information should be combined (e.g., Kahneman & Tversky, 1973):

\[
P(h \mid e) = \frac{P(h)P(e \mid h)}{P(h)P(e \mid h) + P(\neg h)P(e \mid \neg h)}
\]

Bayes’ Theorem prescribes the probability, \(P(h \mid e)\), of experiencing an event \(h\) (e.g., HD) in light of evidence \(e\) (no family history of HD). The best estimate of experiencing that event is a function of the base rate of the event, \(P(h)\), and the diagnosticity of the evidence - the likelihood ratio of true positives, \(P(e \mid h)\), relative to false positives, \(P(e \mid \neg h)\). In Tim’s case, \(P(e \mid h)\) reflects how likely an HD patient is to have no family history of HD, whereas \(P(e \mid \neg h)\) reflects the probability of no family history of HD in those who do not contract HD. From a longitudinal study (Hawe, Talmud, Miller, & Humphries, 2003), we can calculate \(P(e \mid h) = .52\) and \(P(e \mid \neg h) = .66\). Tim’s estimate of \(P(h)\) is 20% and therefore his best estimate of his chance of contracting HD is:

\[
\frac{.2 \times .52}{.2 \times .52 + .8 \times .66} = 16\%
\]

The base rate of HD is actually 33% for 55 year-old males (Bleumink et al., 2004). If Tim receives this information, he should recalculate his personal risk once more, using Bayes’ Theorem, replacing his previous base rate estimate (20%) with 33%, which will result in an increased ‘best estimate’:

\[
\frac{.33 \times .52}{.33 \times .52 + .67 \times .66} = 28\%
\]

Given the two basic components to normative probability judgments – base rates and evidence – there are thus two ways to receive undesirable (desirable) new information: One can receive new, diagnostic information which suggests that one is more (less) at risk, or one may discover that the base rate is higher (lower) than previously thought.

In Equation 2, Tim knows the accurate base rate, calculates his personal risk rationally, and yet his personal risk is different from the base rate. In contrast to the central assumption of the update method, individuals should not necessarily change their estimate of personal risk simply because it lies above or below the base rate. Researchers can only prescribe what effect the new base rate information should have on a participant’s risk estimate if they know the participant’s previous estimate of the base rate. Without this knowledge, it is impossible to classify a particular trial as ‘desirable’ or ‘undesirable’ and therefore to say in which direction (and how much) the participant’s estimate should change. Study 1 highlights the flaws in the update method by simulating optimistic data from non-biased agents!

**Study 1: A Simulation**

Take a hypothetical sample of 100 Bayesian agents, 25 of whom assume base rates of .1, .2, .3, and .4 respectively (mean = .25) for Disease X, which has a true base rate of .25. Before the study, these agents receive evidence
reflecting their vulnerability to Disease X with the following characteristics: $P(e|h) = .5$; $P(\neg e|\neg h) = .9$.

$P(h)P(e|h) + P(\neg h)P(e|\neg h)$ defines the proportion who received evidence suggesting increased risk, here: $0.25 \times 0.5 + 0.75 \times 0.1 = 0.2$. Thus, 20% of agents have evidence suggesting they will get the disease (‘positive evidence’) and 80% have evidence suggesting they will not (‘negative evidence’). At each base rate, 5 agents will receive positive evidence, and 20 will receive negative evidence.

In the study (Table 1), agents calculate their initial risk estimates normatively via Bayes’ Theorem, using their subjective base rates; their second estimate recalculates Bayes’ Theorem using the experimenter-provided true base rate (see Equations 2 and 3).

Agents whose subjective base rate estimates were below the true base rate (.25) truly receive undesirable information: Disease X is more prevalent than they thought. Agents whose subjective base rate estimates were above the true base rate truly receive desirable information: Disease X is less prevalent than they thought. However, the update method classifies information as ‘desirable’ or ‘undesirable’ based on the relationship between initial estimate and true base rate, thus misclassifying 30% of the sample (grey columns).

The final experimental ‘result’ is obtained by averaging across those agents receiving ‘desirable’ and ‘undesirable’ information (as defined by the experimenter). As each positive evidence group represents 5 agents, and each negative group 20 agents, the resulting absolute means are: Desirable Group = 0.04; Undesirable Group = 0.03. Thus, these rational agents show ‘greater updating’ in response to ‘desirable’ than ‘undesirable’ information and would be labelled optimistic – though rational by definition – simply due to incorrect classification. Although this is a somewhat small effect, it becomes much more pronounced when base rate estimates are regressive toward the midpoint of the scale, as is typical of people’s probability estimates in many contexts (see e.g., Harris & Hahn, 2011; Moore & Small, 2008, and references therein). If the true base rate were .21, i.e., below the agents’ mean estimate, then the seeming difference in updating rises to 8% (‘desirable’ = .084; ‘undesirable’ = .004) – easily able to account for extant experimental data. Figure 2 demonstrates that this pattern of results is not dependent on the precise parameters used in this illustrative example. The preponderance of positive differences in updating (where people update more in response to desirable than undesirable information) is clear from both Figure 2a (where mean estimates of the underlying base rate are correct) and Figure 2b (where base rate estimates are regressive, & consequently represents more realistic simulations).

### Implications of the Update Method’s Flaws

The key implication of the simulation results presented above is that a difference in participants’ amount of updating in response to ‘desirable’ and ‘undesirable’ information, as defined in Sharot et al. (2011; see also, Sharot, Guatart-Masip et al., 2012; Sharot, Kanai et al., 2012), cannot be taken as evidence of optimism. As with the statistical confounds in the comparison method, however, there is a simple test that disassociates the predictions of an optimism account from a statistical artifact account: the inclusion of positive events (Harris & Hahn, 2011). Our simulation result is blind to the valence of the events being judged. Consequently, for positive events the statistical account again predicts that updating will be greater in response to information that the base rate is lower than you originally believed your own risk to be. For positive events, which we want to experience, these would be classified as ‘undesirable’ trials, and hence this pattern of results would suggest pessimism. Thus, updating is undoubtedly optimistic (despite the flaws in the method) if participants update more in response to desirable than undesirable information for both negative and positive events.

### Table 1: Artifactual Unrealistic Optimism in Study 1

<table>
<thead>
<tr>
<th>Subjective base rate</th>
<th>Those with positive evidence</th>
<th>Those with negative evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>(n=5)</td>
<td>(n=5)</td>
</tr>
<tr>
<td>Initial Estimate</td>
<td>0.357</td>
<td>0.556</td>
</tr>
<tr>
<td>True Base Rate</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>Experimenter-defined Desirability</td>
<td>Des</td>
<td>Des</td>
</tr>
<tr>
<td>True Desirability*</td>
<td>Undes</td>
<td>Undes</td>
</tr>
<tr>
<td>Correctly Classified?</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>Final Estimate</td>
<td>0.625</td>
<td>0.625</td>
</tr>
<tr>
<td>Amount of update (IE - FE)</td>
<td>-0.268</td>
<td>-0.069</td>
</tr>
</tbody>
</table>

*Des = Desirable information; Undes = Undesirable information
Furthermore, Figure 2 demonstrates how a non-selective change to a probabilistically relevant quantity (e.g., changing the perceived diagnosticity of individuals’ evidence; or altering the regressiveness of their initial base rate estimates) that affects all agents equally, can lead to a seemingly selective effect: a sharp increase in the difference between updating for ‘desirable’ vs. ‘undesirable’ information. Thus, the selective effects of L-DOPA (Sharot, Guitart-Masip et al., 2012) and transcranial magnetic stimulation (TMS) (Sharot, Kanai, et al., 2012) on belief updating might be entirely unrelated to optimism, and simply reflect (for example) better learning (i.e., less conservative updating - formally equivalent to an increase in the diagnosticity of information) following receipt of L-DOPA.

Figure 2: Seemingly ‘optimistic’ updating. Positive values in the z-axes demonstrate ‘optimistic’ updating for Bayesian agents receiving evidence with the properties shown in the x-axes, $P(e|h)$, and y-axes, $P(¬e|¬h)$. Panel a), agents estimate the base rate as either: .1, .2, .3, .4, and the true base rate = .25 (as in Table 1). Panel b), true base rate = .21. The mass of data points in both plots spuriously suggests optimistic updating.

**Study 2: An Experiment**

The results reported above led us to run an experiment in which we followed Sharot et al.’s (e.g., 2011) method, but crucially included a selection of positive events in addition to the negative events. We were further interested in a comparison of ASC participants, with typical adult (TA) controls. Sharot et al. (2011) reported that optimistic updating was largely mediated by the medial prefrontal cortex. Given that ASC is associated with hypoactivity in this region across a range of tasks (Gilbert et al., 2008), and ASC participants have been suggested to be less susceptible to emotional biases in decision making (de Martino et al., 2008), one might predict these individuals to show less of an optimistic bias, similar to the attenuation of the effect in participants who have received TMS to the left inferior frontal gyrus (Sharot, Kanai, et al., 2012). Should this effect be observed, we are crucially interested in the results from the positive events in this study. If TAs do show a seeming pessimism bias (as predicted by the statistical account), will the individuals with ASC show more of a ‘pessimism bias’ (as they appear to be less optimistic than TAs with negative events), or will this effect also be attenuated – suggesting that ASC has not had a selective effect on ‘optimism’?

**Method**

**Participants** 20 male participants with a mean age of 34 years participated in return for £10. The 10 ASC participants all received diagnoses from an independent clinician. Of the 10, using the Autism Diagnostic Observation Schedule (Lord et al., 2000), 5 met the criteria for autism, and 5 for autistic spectrum disorders. The ASC and NT participants were matched for IQ (Weschler, 2011) and age, $t_5 < 1$.

**Design** The experiment was a 2x2x2 (group x event type [positive/negative] x information desirability) design, with event type and information desirability manipulated within-participants. In a departure from Sharot et al.’s (e.g., 2011) method, the base rates presented to participants were derived from participants’ initial estimates of their own risk (IE1), rather than being externally sourced. This enabled the experimental manipulation of information desirability, as well as controlling for the potential differential accuracy of base rate information pertaining to negative events (for which there is plenty of information) versus positive events (for which there is a scarcity of information). Base rates were computed according to the following formula: A random percentage between 17% and 40% (uniform distribution) of IE was either added to, or subtracted from, the IE, according to trial type, and rounded to the nearest integer. Thus, for example, on positive desirable trials a random percentage of the IE was added to the IE resulting in a derived probability indicating that the positive event was more likely to occur than had previously been estimated, and on negative desirable trials a random percentage of the IE was subtracted from the IE, indicating that the negative event was less likely than had been estimated. All
probabilities were capped between 3% and 80% as participants were informed that this was the range of possible probabilities (see also, e.g., Sharot et al., 2011).

Stimuli and Procedure Eighty short descriptions of life events, many of which had previously been used in the study of unrealistic optimism (e.g., Sharot et al., 2011; Weinstein, 1980), were presented in a random order. Half of the events were positive and half negative. Very rare or very common events were avoided and all known probabilities lay between 10% and 70% (M = 32.6, SD = 18.8; Office for National Statistics and PubMed). The procedure followed that of Sharot et al. (2011) (see Figure 1), with one addition. Following their IEs, participants were asked to estimate the chance of the event occurring to the average person. Following the main task, participants’ memory for the base rates presented to them was tested, and they provided four salience ratings for each event. Finally, a funneled debrief procedure (Bargh & Chartrand, 2000) showed that no participants suspected that the derived base rates were inaccurate.

Scoring For each event the amount of update was calculated first by computing the absolute difference between the updated estimate in the second session and the IE, and second, coding the difference as positive when the update was in the direction of the base rate and negative when the update was away from the base rate. Mean updates for each participant in each condition were then calculated after removal of outliers (± 3 x the interquartile range) and trials for which a derived probability could not be applied (e.g., when a participant’s IE was already at the lowest extreme of the probability range, but the trial-type required that a lower base rate be supplied).

Results TA Controls For negative events, we replicated the central result from Sharot et al. (2011). Participants updated more in response to desirable information than to undesirable information, t(9) = 5.02, p = .001 (see Figure 3). This result demonstrates that the minor changes to the update method’s procedure were not consequential for the general paradigm. For positive events, however, the reverse pattern was observed, with greater updating in response to undesirable information than desirable information, t(9) = 4.71, p = .001. Across both event types, a significant event type x desirability interaction was observed, F(1, 9) = 35.45, p < .001, with no main effect of event type or, crucially, information desirability (Fs < 1).

TA Controls and ASC Participants Figure 3 shows that ASC participants showed significantly less asymmetry in their information updating than did TAs, with a significant group x event type x desirability interaction, F(1, 18) =

4.83, p < .041. None of the effects observed in the TAs were observed in the autistic participants (ps > .17). The significant 3-way interaction demonstrated that updating in the ASC group was different from in the TA group, but there was no evidence that they were either more or less optimistic.

Discussion Sharot, Kanai et al. (2012) administered TMS to the left IFG in one group of their participants. Participants subsequently completed an update task comprised entirely of negative events. Participants who had received TMS to the left IFG were shown to exhibit less of an asymmetry in updating to desirable versus undesirable information, and it was therefore concluded that the left IFG is involved in the suppression of updating beliefs about vulnerability in response to undesirable information. Study 2 can be seen as analogous to this study, with the presence or absence of ASC the quasi-experimental variable in place of the location of TMS. Study 2 included positive events in addition to negative events. For positive events, TAs did not show an optimistic bias in belief updating, but rather a seeming pessimistic bias. This resulted in a 2x2 interaction, such that TAs updated more in response to desirable information for negative events, but more in response to undesirable information for positive events. It was this interaction, rather than a specific optimism bias that was attenuated in the ASC participants. One can only speculate whether the same effect would occur with TMS of the left IFG.

Summary Optimism is currently a research area of much interest. Studying optimism with real-world events is complicated by our inability to know the objective probability of an event in a given individual’s future (see also, Weinstein & Klein, 1996). The very prevalent comparative method has been shown to be insufficient (Harris & Hahn, 2011), and here
we have demonstrated that a recent (‘update’) method also fails as a test of human rationality, through ‘demonstrating’ that rational agents are irrational (Study 1)! Despite this weakness, the addition of positive events can determine whether there is evidence for an optimistic bias. With this addition, no evidence is found (Study 2). The inclusion of ASC participants in Study 2 also demonstrated the difficulties inherent in making conclusions about the effects of individual differences or experimental manipulations on optimistic belief updating, without the inclusion of both positive and negative events.

Our resulting conclusion, that there is no evidence for a general optimistic bias in human likelihood judgments, might seem surprising in the light of historical common belief. It is, however, consonant with recent results where participants overestimate the likelihood of negative events relative to neutral events in the laboratory (e.g., Bilgin, 2012; Harris, Corner, & Hahn, 2009; Risen & Gilovich, 2007; Vosgerau, 2010). On the basis of the extant evidence, we cannot conclude that an optimistic view of the future characterises healthy human thought (Taylor & Brown, 1988).

References