Why the Brain Talks to Itself: Sources of Error in Emotional Prediction

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citation

Published Version
http://dx.doi.org/10.1098/rstb.2008.0305

Citable link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:3110936

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
Why the brain talks to itself: sources of error in emotional prediction

Daniel T. Gilbert and Timothy D. Wilson

Phil. Trans. R. Soc. B 2009 364, 1335-1341
Mark Twain worked hard to be funny. One of his editors reported that before every speech, Twain ‘mused his words to an imagined audience ... He studied every tone and every gesture and he forecast the result with the real audience from its result with that imagined audience’ (Twain 1910, p. 11). Twain tested new material on an imaginary focus group and assumed that if the people in his head laughed at a punch line then people in the theatre would do the same. This technique may seem unremarkable, but in fact, there is something decidedly odd about it. After all, Mark Twain was testing jokes on the people in his head, but the people in his head were all ... well, Mark Twain. That is, the person who was telling the joke and the people who were reacting to the joke were all inventions of the same brain—so why did that brain need to go through an elaborate ritual of telling and listening just to find out whether its own jokes were funny?

If the brain were a single unified system, then this ritual would be puzzling. But the brain is not a single unified system. Rather, it is a collection of independent systems that specialize in receiving, processing, producing and transmitting different kinds of information. The parts of Mark Twain’s brain that produced jokes were not same parts that produced laughter, so to determine whether a joke was funny, the joke-production system had to tell the joke to the laughter-production system and then take note of its reaction. As it turns out, this ritual is the one that most of us perform every day—not just to find out whether our jokes are funny, but to make the predictions by which we guide our lives.

1. FROM PREVIEW TO PREMOTION TO PREDICTION
Organisms remember the past so that they can predict the future. If a bird can remember the appearance of a cat followed by the tinkling of a bell, the bird can thereafter use the sound of the bell to predict the cat’s arrival. This is a valuable skill because a bird that can predict the future can take action to preclude those futures in which it plays the role of the entree. The memory of the co-occurrence of the cat and the bell allows the bird to transcend the normal restrictions of linear time, reacting to a future cat before it becomes a present cat. The brain specializes in memory because memory enables prediction, and prediction gives organisms a head start. Alas, memory-based prediction requires past experience: a bird can predict the appearance of a cat from the sound of a bell only if those two stimuli have co-occurred in the past, which means that a bird can act pre-emptively only after having had at least one close encounter of the feline kind. Because such encounters are a potentially expensive way to gain the power of prognostication, human beings have developed a different and more sophisticated technique that allows them to make predictions about future events they have never experienced before.

For example, when asked how much they would enjoy learning that they had pancreatic cancer, most people can reliably produce the correct answer, which is some version of ‘not so much’. They do not produce this answer by remembering how it felt to receive such news in the past, but by closing their eyes for a moment, simulating the event and then noting their emotional reaction to that simulation. Just as Mark Twain assumed that feeling amused by an imaginary joke meant that he would be amused by the real thing, most of us assume that if we feel anxious when we imagine bad news then we would feel even more anxious receiving it. In essence, we generate mental simulations or previews of future events, which cause us to have affective reactions or premotions, which we then use as a basis for our forecasts or predictions about the event’s likely emotional consequences (Damasio 1994; Schwarz & Strack 1999; Breiter et al. 2001; Berns et al. 2006; Gilbert 2006; Gilbert & Wilson 2007). We know which future events will feel good and which will feel bad because we feel good or bad when we simulate them.

Why the brain talks to itself: sources of error in emotional prediction

Daniel T. Gilbert1,* and Timothy D. Wilson2

1Department of Psychology, Harvard University, Cambridge, MA 02138, USA
2Department of Psychology, University of Virginia, Charlottesville, VA 22904, USA

People typically choose pleasure over pain. But how do they know which of these their choices will entail? The brain generates mental simulations (previews) of future events, which produce affective reactions (premotions), which are then used as a basis for forecasts (predictions) about the future event’s emotional consequences. Research shows that this process leads to systematic errors of prediction. We review evidence indicating that these errors can be traced to five sources.

Keywords: emotional prediction; affective forecasting; prediction
2. ERRORS IN EMOTIONAL PREDICTION

Previews and premotions are the building blocks of prediction, and neuroscientists have recently begun to investigate them. For example, research suggests that previews of future events are produced in large part by the frontal regions of the brain (Ingvar 1985; Wheeler et al. 1997; Fellows & Farah 2005), which are especially activated when people are asked to simulate the future (Addis et al. 2007; Buckner & Carroll 2007; Schacter et al. 2007; Szpunar et al. 2007). Patients with damage to these regions are often unable to simulate future events (Tulving 1985; Tulving et al. 1988; Klein et al. 2002) and are typically described as being ‘bound to present stimuli’ (Melges 1990) and ‘locked into immediate space and time’ (Faglioni 1999). The premotions that these previews produce appear to depend specifically on the ventromedial prefrontal cortex, as people with damage to this area find it difficult to predict their emotional reactions to the events they are previewing (Bechara & Damasio 2005; Shiv et al. 2005).

While neuroscientists have been investigating the neural substrates of previews and premotions, psychologists have been investigating how and how well they allow people to predict their emotional reactions to future events (for recent reviews see Loewenstein & Angner 2003, Wilson & Gilbert 2003). This research has uncovered two general sources of error. When we generate a preview of a future event, our premotions are influenced by the content of that preview (i.e. those features of the future event that appear in our mental simulation) and by the context of that preview (i.e. the features of the current situation in which we are generating the mental simulation). When the event actually happens, our emotions are influenced both by the content and the context of our view. It follows, then, that premotions will be good predictors of emotions when the content and context of the event we preview are similar to the content and context of the event we view. For example, if one sunny day a man was on his way to his favourite cafe and imagined eating the chocolate cake, he would probably experience a positive premotion and expect to enjoy eating the cake when he arrived. If he arrived while the sun was still shining and ate the chocolate cake he had imagined, there is every reason to believe that he would enjoy the experience just as he predicted (figure 1). On the other hand, if he arrived at the cafe only to find that (a) the chocolate cake was gone and mincemeat pie was being served in its place (figure 2) or (b) the weather had turned cold and nasty (figure 3), then he might not enjoy his experience as much as he expected to. The fact is that premotions accurately predict emotions when the content and context of the event we preview are similar to the content and context of the event we view. For example, if one sunny day a man was on his way to his favourite cafe and imagined eating the chocolate cake, he would probably experience a positive premotion and expect to enjoy eating the cake when he arrived. If he arrived while the sun was still shining and ate the chocolate cake he had imagined, there is every reason to believe that he would enjoy the experience just as he predicted (figure 1). On the other hand, if he arrived at the cafe only to find that (a) the chocolate cake was gone and mincemeat pie was being served in its place (figure 2) or (b) the weather had turned cold and nasty (figure 3), then he might not enjoy his experience as much as he expected to. The fact is that premotions accurately predict emotions when the content and context of the event we preview are similar to the content and context of the event we view, and the reason why errors in emotional prediction occur is that these two criteria often go unmet. We shall now explain why.

(a) The problem of dissimilar content

Events do not always unfold precisely as we imagine them. One especially dull reason for this is that the future is inherently uncertain: no matter how hard we try, we can never know everything there is to know about the cakes and cafes that populate our tomorrows, and even our best guesses are nothing more than that. The inherent uncertainty of the future means that previews and views often have dissimilar content. Research shows that several other decidedly less dull factors exacerbate these differences.

(i) Previews are unrepresentative

People who have difficulty remembering the past often have difficulty predicting the future (Tulving 1985; Tulving et al. 1988; Klein et al. 2002).
Klein et al. 2002; Hassabis et al. 2007). When we generate previews of a future event, we draw on a vast network of information about similar events that have happened in the past (Hawkins & Blakeslee 2004; Dudai & Carruthers 2005; Addis et al. 2007; Buckner & Carroll 2007), and thus our previews are only as good as the memories on which they are based. Ideally, a preview of a future event should be based on memories of very similar past events—but given that we cannot know precisely how a future event will unfold, how can we know which past events are the most similar to it? The statistically sensible solution to this problem is to base our previews on those past events that are most representative or typical of their class. For instance, we may not know precisely what our next dental appointment will be like, but we do know that most past dental appointments have involved cleaning and not extraction, and thus our preview of the next visit has a better chance of being accurate if it is based on memories of flossing and buffing rather than on memories of injecting and drilling.

But research shows that we tend to base our previews on those memories that are most available rather than the most typical and that, ironically enough, our most available memories are often of the least typical events. For example, in one study, commuters who were waiting on a platform were asked to imagine how they would feel if they missed their train that day (Morewedge et al. 2005). Before making this prediction, some of the commuters (the any-memory group) were asked to remember and describe ‘a time you missed your train’. Other commuters (the worst-memory group) were asked to remember and describe ‘the worst time you missed your train’. The results showed that commuters in the any-memory group remembered a past episode that was every bit as bad as the episode remembered by commuters in the worst-memory group. Apparently, disastrous but atypical instances of train missing came more readily to mind than did less disastrous but more typical instances. As such, when commuters in the any-memory group were asked to predict how they would feel if they missed their train that day, they expected to feel terrible—and that prediction turned out to be wrong. Commuters based their predictions on the premonitions that their previews produced, but because these previews were based on atypical memories, they provided a poor basis for prediction.

Like atypical experiences, recent experiences are especially available and are thus likely to become the building blocks of our previews. Participants in one study (Kahneman et al. 1993) were asked to submerge their hands in ice water while using an electronic rating scale to report their moment-to-moment discomfort. All participants performed a short trial and a long trial. On the short trial, participants submerged their hands for 60 s in a water bath that was 57°F. On the long trial, participants submerged their hands for 90 s in a water bath that was 57°F for the first 60 s and then surreptitiously warmed to 59°F over the course of the remaining 30 s. The short trial thus consisted of 60 cold seconds, and the long trial consisted of the same 60 cold seconds with an additional 30 less-cold seconds. Although the participants’ moment-to-moment reports revealed that they experienced equal discomfort for the first 60 s on both trials but much more discomfort in the next 30 s on the long trials, they remembered the long trial as less painful than the short trials because it had a less painful ending. When participants were then asked which of the two trials they would prefer to repeat (a question that required them to preview those future events and predict how painful they would be), 69 per cent chose to repeat the long trial.

The first reason why previews provide a poor basis for prediction, then, is that they tend to be based on memories that are not representative of the future events that those previews were meant to simulate.

Previews are essentialized

If previews contained every detail of the views they were meant to simulate, then imagining a dental appointment would take precisely as long as the appointment itself. But it does not, and that is because previews generally contain only the essential features that define an event and omit the features that are merely incidental to it. When we preview a dental appointment, we imagine sitting in the dentist’s chair or having our teeth inspected but not parking the car or leafing through magazines in the waiting room. The benefit of omitting these incidental features from our previews is that we can preview a 30 min appointment in 15 s. The cost of omitting these incidental features is that although they do not define the event, they can have a significant influence on our emotional reactions to it.

Most events have a small set of extremely positive or extremely negative features that define them and a much larger set of mildly positive and mildly negative features that are incidental to them. Having someone work on our teeth is the defining feature of a dental appointment, and for most of us this feature is quite negative. On the other hand, parking the car and reading magazines are incidental features of the event and these features may be mildly positive (free parking) or mildly negative (old magazines). Our emotional reaction to the actual dental appointment will be a weighted sum of its defining and incidental features. The defining features are likely to be more powerful than the incidental features, of course, but the fact that there are so many incidental features means that they too may have a considerable combined impact. Because some of these incidental features are likely to have a valence opposite to that of the defining features, these incidental features are likely to dilute the power of the defining features. In other words, we imagine the unpleasant features of going to the dentist because these features are what going to the dentist is all about, and we fail to imagine the pleasant features of going to the dentist because they are incidental to the purpose of our visit. Anyone who has ever expected to enjoy a gourmet meal at a famous restaurant and instead spent the evening being irritated by the interminable wait and the imperious waiter has experienced this phenomenon. Blanquette de veau may define a gastronomic experience and waits and waiters may be incidental to it, but the latter features are likely to have some impact on one’s experience of the event.

Studies confirm that our failure to preview the incidental features of future events can lead us to mispredict our emotional reactions to them.
For example, participants in one study (Wilson et al. 2000) were asked to predict how they would feel the day after their favourite football team won or lost a game against its arch rival. Before making these predictions, some participants were asked to describe the day’s activities in great detail, i.e. to preview the incidental features of the event. Other participants were not asked to preview the incidental features. The results showed that the participants who were not asked to preview the incidental features of the event expected to be very happy if their team won and very unhappy if their team lost, but that those who were instructed to preview the incidental features of the event made much more moderate emotional predictions—and these more moderate predictions turned out to be more accurate. Similar studies have shown that people overestimate how happy they would be if they moved to California (Schkade & Kahneman 1998) or became wealthy (Kahneman et al. 2006) because their previews of these events include defining features such as sunshine and money, but fail to include incidental features such as traffic and taxes.

Our tendency to omit incidental features from our previews of future events is especially pronounced when the events are temporally distant (Vallacher & Wegner 1985; Trope & Liberman 2003). Participants in one study (Liberman & Trope 1998) were told that in a year there would be a lecture on an interesting topic at an inconvenient location and a lecture on a boring topic at a convenient location. Participants were asked to predict which lecture they would attend. Because their previews contained the defining feature of the event (e.g. the topic) but lacked the incidental feature (e.g. the location), participants predicted that they would attend the more interesting lecture, presumably because they expected to enjoy that experience more than its alternative. But participants who were told that the same lecture was taking place tomorrow instead of next year were more likely to include the incidental feature of the event in their previews, and were thus more likely to predict that they would attend the lecture that was boring but convenient.

The second reason why previews and views differ, then, is that previews tend to omit features that are incidental to the event but that nonetheless may have a significant impact on our emotional reactions to it.

(iii) Previews are truncated

Just as previews tend to emphasize the defining rather than the incidental features of future events, so do they tend to emphasize the event’s early occurring rather than late-occurring moments. For example, in one study (E. W. Dunn & C. E. Ashton-James 2008, unpublished data), people were asked to predict how much they would enjoy an experience that involved putting their hand in an ice water bath for a few minutes (which is painful) and receiving a massage (which is pleasant). Participants in the good beginning condition were asked to predict how much they would like the experience if the massage preceded the ice water bath, and participants in the bad beginning condition were asked to predict how much they would like the experience if the ice water bath preceded the massage. Results showed that participants in the good beginning condition expected to like the experience more than did participants in the bad beginning condition when, in fact, the order of the two components had no influence on how much participants actually liked the experience. Just as people represent past events by their endings, they appear to represent future events by their beginnings.

One important consequence of emphasizing early occurring moments is that previews take little account of adaptation. For many reasons, emotions tend to dissipate over time, which means that previews tend to emphasize precisely those moments that evoke the most intense emotion. This leads to one of the most pervasive errors of emotional prediction—the impact bias—which is the tendency for predicted emotions to be more extreme than actual emotions (Wilson & Gilbert 2003). For example, when people are asked to imagine how they would feel some time after a negative event (such as failing to get tenure or breaking up with a romantic partner) or a positive event (such as getting tenure or beginning a new romantic relationship) they typically overestimate how bad or good they will feel because they fail to simulate their adaptation to the event (Gilbert et al. 1998). Similarly, research shows that healthy people consistently overestimate how unhappy they would be in various states of ill-health (Menzel et al. 2002; Ubel et al. 2003; Riis et al. 2005) because their previews of illness emphasize the early moments in which they are making the difficult transition from being healthy to being ill, but not the more numerous moments that follow, in which they have adapted to their new state. People imagine becoming disabled rather than being disabled (Kahneman et al. 2006), and becoming is much worse than being because over time most people adapt at least partially to disability (Lucas 2007; Oswald & Powdthavee 2008).

Because previews tend to ignore adaptation, they tend to be insensitive to those features of an event that might promote or inhibit it (Gilbert et al. 1998, 2004a,b; Gilbert & Ebert 2002). For example, research has established that people adapt to events more quickly when they understand why those events happened than when they do not (Wilson & Gilbert 2008). When participants in one study (Wilson et al. 2005) were approached at a library and given $1, those who received an explanation for the gift were less happy a few minutes later than were those who received no explanation. But when asked to preview these two incidents, a similar group of participants predicted that they would be happier if they received the gift with an explanation rather than without one. Because their previews did not include the event’s late-occurring moments, people who were asked to preview receiving a gift with an explanation were unable to foresee their own adaptation. Similarly, research has established that people adapt to events more quickly when they cannot undo them than when they can. Participants in one study (Gilbert & Ebert 2002) were more satisfied with a gift when they did not have the opportunity to exchange it than when they did; and yet, a similar group of participants who were asked to preview these two incidents were insensitive to the opportunity for exchange and expected to be just as happy in both instances. Because their previews did
not include the event’s late-occurring moments, participants who were asked to preview receiving a gift that they could not exchange were unable to foresee their own adaptation.

The third reason why previews and views differ, then, is that they tend to emphasize the early occurring moments of the event in which emotions are likely to be the most intense.

(iv) Previews are comparative
How would it feel to buy a lottery ticket that paid $50 if one’s friend bought a ticket that paid $80? Many of us have the compelling intuition that we would be slightly unhappy, and that we might actually be happier if we had won only $40 and our friend had won only $10 (Tversky & Griffin 1991; Solnick & Hemenway 1998). The reason we make this prediction is that we imagine comparing our $60 to our friend’s $80, which makes our winnings seem paltry by contrast. But research suggests that in a wide range of circumstances, people are less likely to make such comparisons than they imagine (Novemsky & Ratner 2003; Hsee & Zhang 2004). For example, in one study (Morewedge et al. 2007), participants were either asked to eat or to imagine eating a potato chip in the presence of a superior food (chocolate) or an inferior food (sardines).

Participants who imagined eating the chip predicted that they would enjoy it more when it was eaten in the presence of sardines than in the presence of chocolate. In other words, these participants mentally compared the food they would be eating with the food they would not be eating. But participants who actually ate the chip enjoyed it just as much when it was eaten in presence of chocolate as when it was eaten in presence of sardines. Similarly, students in another study (Golub et al. in press) who imagined receiving a poor grade on their midterm exam predicted that they would feel worse if they had been expecting a good grade than if they had been expecting a bad grade. In other words, they believed that they would compare the grade they received to the grade they had expected to receive. In actuality, however, students who received a poor grade felt equally bad regardless of their prior expectations. In both of these studies, people’s previews included comparisons that their views did not.

Just as our previews omit incidental features that ultimately impact our emotional reactions, they tend to include comparative features that ultimately do not impact our emotional reactions. Although we think we will compare what we got with what others got, with what we thought we would get, with what we could have got, or with what we have got in the past, the fact is that real outcomes command our attention and thus attenuate such comparisons. When we imagine eating chips in the presence of chocolate, we naturally compare the two; but when we actually have a mouthful of crunchy, salty, greasy, fried potatoes, the experience we are having is much more salient than the one we are not having, which makes comparison less likely. Imaginary chips are readily compared to imaginary sardines, but real chips are not. The fourth reason why previews and views differ, then, is that previews include comparisons that views do not.

(b) The problem of dissimilar context
Accurate predictions require that the content of our previews be similar to the content of our views, and as the studies reviewed above suggest, this is not always the case. But accurate predictions also require that the context in which previewing occurs be similar to the context in which viewing occurs, and as it turns out, this is not always the case either. Why do contexts matter? Premotions are not just reactions to previews; they are reactions to previews and to the context in which those previews are generated. That is why we feel happier when we preview chocolate cake while we are lying on a comfortable couch than on a bed of nails, or when we are hungry rather than sated. When viewing immediately follows previewing—for example, when we see a doughnut, buy it and pop it into our mouths—the contexts in which these two operations were carried out are likely to be similar. But when previewing precedes viewing by a substantial interval—for example, when we see a doughnut, buy it, take it home and eat it for breakfast the next morning—the two contexts are likely to differ. When this happens, the premotions we experienced at the bakery may be unreliable indicators of the emotions we will experience when we eat the doughnut at home the next day.

All of this seems elementary, but research shows that people often fail to realize just how easily differences in the contexts of previewing and viewing can derail emotional prediction (Loewenstein et al. 2003). For example, in one study (Gilbert et al. 2002) hungry and sated people were asked to predict how much they would enjoy eating a bite of spaghetti with meat sauce for dinner the next day or for breakfast the next day. The results showed that the people’s current level of hunger strongly influenced their predictions and that the time of day at which they would be eating the spaghetti did not. In other words, hungry people mistakenly expected to like eating spaghetti for breakfast the next day and sated people mistakenly expected to dislike eating spaghetti for dinner the next day. In a related study (Van Boven & Loewenstein 2003), people who were working out at a gym were asked to predict how much they would want food and water if they were lost in the woods. Those who had just finished exercising predicted that they would want water more than food, but those who were just about to exercise predicted that they would want food more than water. In both of these studies, people failed to realize that the context in which they were generating their previews was having a strong influence on their premotions, and that because the previewing context was different from the viewing context, their premotions would not match their emotions.

3. CODA
Emotional prediction is the process by which we discover what we already know. We evolved to have emotional reactions to events in the present, and thus, to find out how we will react to events in the future, we simply pretend those events are happening now. But the system that does the reacting (often called system 1) and the system that does the pretending (often called system 2) are not the same systems (Stanovich 1999).
System 2 generates simulations of dentists and doughnuts and system 1 generates emotional reactions to those simulations. Because system 2 does not have direct access to the information that system 1 uses to generate these emotional reactions, it tells system 1 a fairy tale—sometimes enthralling, sometimes frightening—and then listens closely to its response. System 2 uses system 1 as a test audience and assumes that if system 1 reacts with pleasure to a simulated event, then it will react with pleasure to the event itself. In essence, system 2 determines how system 1 will react in the future by tricking it into reacting in the present.

This ritual of pretending and reacting, telling and listening is an inventive way to make emotional predictions, but it suffers from two shortcomings. First, the simulations that system 2 generates are not faithful facsimiles of the events they are meant to model. Compared to real events, simulated events are unrepresentative (i.e. they are based on past events that are not typical of their class), essentialized (i.e. they omit features that are incidental to the event but that influence emotional reactions to it), truncated (i.e. they emphasize early occurring moments in which emotions are most intense, and ignore late-occurring moments in which adaptation occurs) and comparative (i.e. they include comparisons that are not unlikely to be made during viewing). Second, both the premotional and emotional reactions that system 1 generates are influenced by the contexts in which they occur, and these contexts are not necessarily the same. For both of these reasons, our premotional reactions to simulated events often differ from our emotional reactions to the events themselves, rendering our predictions inaccurate. When the human brain talks to itself, it does not always tell the truth.

We gratefully acknowledge the support of research grant no. BCS-0722132 from the National Science Foundation to D.T.G. and T.D.W. We thank Lisa Feldman-Barrett for comments on an earlier version of this paper.

REFERENCES


Phil. Trans. R. Soc. B (2009)


