Understanding Surprise: Can Less Likely Events Be Less Surprising?

Rebecca Maguire (rebecca.maguire@dbs.edu)
Department of Psychology, Dublin Business School
Dublin 2, Ireland.

Phil Maguire (pmaguire@cs.nuim.ie)
Department of Computer Science, NUI Maynooth
Co. Kildare, Ireland.

Mark T. Keane (mark.keane@ucd.ie)
School of Computer Science and Informatics, University College Dublin
Belfield, Dublin 4, Ireland.

Abstract

Surprise is often thought of as an experience that is elicited following an *unexpected* event. However, it may also be the case that surprise stems from an event that is simply difficult to *explain*. In this paper, we investigate the latter view. Specifically, we question why the provision of an enabling factor can mitigate perceived surprise for an unexpected event despite lowering the overall probability of that event. One possibility is that surprise occurs when a person cannot rationalise an outcome event in the context of the scenario representation. A second possibility is that people can generate plausible explanations for unexpected events but that surprise is experienced when those explanations are uncertain. We explored these hypotheses in an experiment where a first group of participants rated surprise for a number of scenario outcomes and a second group rated surprise after generating a plausible explanation for those outcomes. Finally, a third group of participants rated surprise for the both the original outcomes *and* the reasons generated for those outcomes by the second group. Our results suggest that people *can* come up with plausible explanations for unexpected events but that surprise results when these explanations are uncertain.

Keywords: Surprise, expectation, representation, discourse.

Introduction

In day-to-day life, people have a remarkable ability to make sense of their surroundings and can effortlessly infer connections between events in order to create a rich and detailed representation of any given situation. Nevertheless, this coherent representation of the world can sometimes break down. More specifically, it is known that certain events have the potential to *surprise* us. Far from being an isolated occurrence, surprise is actually quite a common experience. Because of its prevalence, this phenomenon has received a great degree of research attention in cognitive science and psychology. Historically, Darwin (1879) was the first to classify it as one of the most basic emotions, a claim that has been adopted by many subsequent theorists. As well as being associated with a distinct subjective and physiological response, surprise is known to have some important cognitive manifestations. For example, the perception of a surprising event will usually cause a person to cease what they are currently doing and focus their attention on the event in question (Meyer, Reisenzein & Schützwohl, 1997). The purpose of such a reaction is to discover *why* the surprising event transpired, so that a similar event can be anticipated in future circumstances.

In this paper, we explore why people find certain events surprising and other events unsurprising. Specifically, we investigate why presenting an enabling condition for an unexpected event mitigates the level of surprise elicited by that event.

Surprise as unexpectedness

The most intuitive way of describing a surprising event is to say that it was *unexpected*. Likewise, it makes sense to assume that any expected event would be unsurprising if it were to occur. However, this account can be problematic, mainly due to the disagreement surrounding what it means to expect something. For instance, if we relate expectations to probabilities, then every low probability event should be extremely surprising, and vice versa. Evidently however, this is not always the case. For instance, while the outcome of a lottery draw always has an extremely low probability, it is rarely surprising

In light of this, Teigen and Keren (2003) suggested that surprise at a given event might be more accurately explained in terms of its subsequent comparison with an alternative outcome. Investigating this hypothesis, they carried out an experiment which described Erik, an athlete competing in a 5,000m race. In one condition, participants were informed that Erik was in second place behind a lead runner, while in another condition they were told that all the athletes, including Erik, had formed a large group as they approached the finish line. When asked to indicate how surprised they would be if Erik won the race, participants in the first condition (where Erik was in second place) gave slightly higher surprise ratings than those in the second condition (where all the athletes had formed one group), despite the fact that participants correctly rated Erik's *probability* of winning the race as higher in the first condition. One explanation for this result is that the first scenario induces an expectation (that the lead runner will win the race) which is disconfirmed by Erik winning. On the other hand, when all the athletes have an equal chance of winning the race, no expectation is contradicted if Erik wins. This finding can be said to support an *Expectation-Disconfirmation* hypothesis of surprise (see also Meyer et al, 1997).

Despite the intuitive plausibility of this hypothesis however, it may not always be the case that disconfirmed expectations lead to such a high level of surprise. Maguire and Keane (2006) proposed that surprise may be better thought of in terms of *Representation-Fit.* They pointed out that while disconfirmed expectations may frequently lead to perceived surprise, this may not always be the case. For instance, if a person can account for *why* an expectation was disconfirmed, then they might not be so surprised by it.

Surprise as representation-fit

Zwaan and Radvansky (1998) have shown that during reading, people routinely construct situation models, or rich representations, of the depicted events in a discourse. These consist of a number of complex inferences about the central characters, their goals and actions, as well as more general information about the story's temporal and spatial context. As the reader encounters new events, this representation must be continually updated, a process motivated by the need on the part of the reader to achieve coherence among the text constituents (Graesser, Singer & Trabasso, 1994). Accordingly, each new event in a text must be coherently integrated into the existing discourse representation for successful comprehension to result.

Based on this premise, Grimes-Maguire and Keane (2005b) devised a theory of representation-fit for surprise (see also Maguire & Keane, 2006). In short, this theory predicts that the more difficult it is for an individual to coherently integrate a new event into their discourse representation, the more surprising that event will appear. As well as being an intuitive view, the underlying principles of this theory rest on many well supported models of comprehension (e.g. *Constructivist theory*, Graesser et al, 1994; *Landscape model,* Linderholm, Virtue, van den Broek & Tzeng, 2004; *Situation models,* Zwaan & Radvansky, 1998). The main way in which this account differs from existing theories of surprise is that it does not view the process as being dependent on expectation. Instead surprise is considered as a retrospective judgement relating to how well a given event can be connected with those that have preceded it, like trying to fit a piece into a jigsaw puzzle. Consider a scenario, for instance, where you walk out your front door only to find that your car is no longer in the driveway. This is obviously an unexpected, or schema-discrepant, event. However, if this triggers your memory that the car is currently being serviced, you will no longer be surprised since a satisfactory explanation for the unexpected event has been identified. Conversely, there are other situations where you might not experience any surprise until the point where you realise that an event cannot be easily explained. Consider, for example, meeting someone on the street but only later realising that they were supposed to be away on holiday.

A study conducted by Grimes-Maguire and Keane (2005a) offered substantive empirical evidence for these ideas. They found that when participants were asked to indicate their level of surprise for the end event in a scenario, they were extremely adept at detecting subtle differences in how strongly that event could be supported by prior events. They also observed that surprise ratings were *not* correlated with on-line expectations, or forward inferences, arguing against the claim that these two variables are linearly related.

Scenario body	
Anna has a very important job interview in the morning. She has to get up far earlier than usual, so she makes sure to set her alarm clock radio for 7am.	
	How surprised would you be if….?
1	The alarm clock woke her up at 7am ***(Expectation Confirmed)***
2	The alarm clock failed to ring at 7am ***(Expectation Disconfirmed)***
3	There was a power-cut during the night and the alarm clock failed to ring at 7am ***(Expectation Disconfirmed+Enabling Event)***
4	She had a quiet, good night's sleep and the alarm clock failed to ring at 7am ***(Expectation Disconfirmed+Control event)***

Table 1: Sample scenario alongside four experimental conditions examined by Maguire and Keane (2006)

More recently, Maguire and Keane (2006) explored whether Representation-Fit is a better explanation for perceived surprise than the Expectation-Disconfirmation hypothesis. They presented participants with simple everyday scenarios, such as that in Table 1, and asked them to rate surprise for one of four hypothetical endings. In the first condition, the ending was directly in line with the expected outcome. In the second condition, this ending disconfirmed the expectation (e.g. the alarm clock failing to ring at 7am goes against the content of the scenario body). In the critical third condition, participants were asked to rate their surprise for the same unexpected ending as in the preceding condition, alongside a potential enabling factor for that event. The fourth condition acted as a control whereby the same unexpected event was coupled with an irrelevant event that bore no causal relationship to it.

The resulting surprise ratings revealed that participants were significantly *less* surprised by the events in the third condition (i.e. 'Expectation Disconfirmed with Enabling Event') than in the second condition (i.e. 'Expectation Disconfirmed' alone). This suggests that surprise ratings were based on the ease with which the events could be connected with the previous scenario representation, rather than on the mere unexpectedness or probability of those events. Indeed, events in the third condition were *less* probable than those in the second condition, as a conjunction of two events is always logically less likely than one of those events on its own. Subsequent experiments ruled out the possibility that participants were interpreting the enabling condition as a 'given' in the scenarios (Maguire & Keane, 2006). Also, interestingly, the enabling events in isolation were rated as more surprising than when they formed part of the conjunction.

These results support the Theory of Representation-Fit. However, while they demonstrate that the provision of an enabling condition lowers the surprise for a scenario, it is not clear *why* this should be the case. In the following experiment we investigate this matter in more detail.

Experiment

One intuitive explanation for Maguire and Keane's (2006) findings is that participants became surprised in the Expectation Disconfirmed condition because they did not generate a plausible explanation for the unexpected event (e.g. they could not understand *why* the alarm clock failed to ring at 7am). According to this view, the Expectation Disconfirmed + Enabling Event condition appeared less surprising because an explanation was suggested (e.g. a power-cut during the night), thereby offering participants a means of integrating the unexpected event into their representation. If this was indeed the reason for the observed effect, it is important to establish whether participants were actually incapable of explaining the events in question or whether they simply did not generate such inferences spontaneously. An alternative and more intriguing possibility is that participants *were* able to generate plausible explanations, but that this did not mitigate the overall level of surprise. Surprise instead could be due to the *uncertainty* of the actual explanation.

In the following experiment, we differentiate between these two conflicting hypotheses using three between-participant conditions. Firstly, a *Control* group of participants were asked to rate surprise for the Expectation Disconfirmed (hereafter D) and Expectation Confirmed (hereafter C) scenarios used by Maguire and Keane (2006). A second *Generative* group carried out the same task but were first asked to generate a plausible reason for why they thought these events occurred (e.g. *"why do you think the alarm clock failed to ring at 7am?"*). A third *Conjunction* group of participants were asked to indicate how surprised they would be by the occurrence of the same events *in conjunction* with the reasons generated by the second group.

If participants cannot generate convincing explanations for the D scenarios, then we would expect no difference in surprise ratings between the Control and Generative groups, but higher ratings for the Conjunction group (reflecting the unsatisfactory reasons generated). However, if participants do not spontaneously generate enabling conditions for unexpected events but are able to do so when explicitly requested, then surprise ratings should be lower for the Generative and Conjunction groups relative to the Control group. Another possibility is that a greater level of surprise is elicited when the enabling condition is *uncertain*. Thus, for example, participants may hypothesise that a power-failure caused the alarm clock to stop working, but the outcome event may still seem surprising because they cannot be certain of this explanation. If this is the case, then participants in the Conjunction group should give lower surprise ratings relative to the Control and Generative groups, since they are provided with enabling conditions as part of the outcome event, while participants in the other groups are required to hypothesise the enabling conditions.

Method

Participants A total of 100 undergraduate students from UCD took part in this experiment for partial course credit. Data from five participants were discarded due to a failure to complete the experiment.

Materials The same 16 scenarios as used by Maguire and Keane (2006) were employed for this experiment. Only the conditions of C (Expectation confirming scenarios) and D (Expectation disconfirming scenarios) were examined. For the Conjunction group, these conditions were paired with participant-generated enabling events, as described below.

Design The experiment had two stages. In the first stage, one group of participants (the *Control group*) were asked to read each scenario and indicate how surprising they found the final event, while another group of participants (the *Generative group*) were asked to write an explanation for this final event before rating it for surprise. Each participant in these two groups was randomly assigned to read six D and six C scenarios. The second stage of the experiment involved one group of participants (the *Conjunction group*) being assigned to read the same D and C scenarios as the other two groups, along with the dominant enabling event generated by the Generative group. In sum, this was a 3 (Group) x 2 (Scenario Type) design, where Group was a between-participants factor and Scenario Type was a within-participants factor.

Procedure For the first stage of the experiment, participants were randomly assigned to one of the two experimental groups as outlined above (Control or Generative). Participants in the Control group were told that they would be presented with a number of short stories and asked to indicate how surprising they found the final event on a scale of 1 – 7. Participants in the Generative group were also told that they would be presented with a number of stories, but were asked to write down a plausible reason for *why* the final event in the story occurred. Thus, for example, in the D scenario above, they were asked to indicate why they thought the alarm clock failed to ring at 7am. Following this, participants indicated how surprising they found this final event (as opposed to how surprising they found their generated explanation for the event). Each scenario was presented on a separate page and in a different random order for each participant.

Prior to the second stage of the experiment, the responses from the Generative group were analysed and the dominant rationalisation for each of the scenarios was identified. In order to do this, each participant's response to each scenario was categorised in terms of a distinct theme and the most common of these was identified (e.g. for the alarm clock scenario, responses fell into the theme of "battery-failure" or of "setting the alarm clock incorrectly"). It should be noted that there was great uniformity in the reasons generated. That is, the majority of participants generated the *same* potential reasons for the unexpected events. Inter-rater reliability for the categorization process was high (above 95%) and any disagreement was resolved by discussion. The dominant enabling event for each scenario was then presented alongside the original D or C outcome. So, for example, the most commonly generated response for the disconfirming scenario in Table 1 was *"the batteries in the alarm clock ran out"* and this was added to the original outcome to become *"The batteries in the alarm clock ran out and the alarm clock failed to ring at 7am"*. The Conjunction group were then presented with the original scenarios and asked to give surprise ratings for the conjoined outcomes.

Results and Discussion

While there was no difference in the surprise ratings between the Generative and the Control groups, the Conjunction group gave significantly lower surprise ratings for the scenario outcomes. These results can be seen in Figure 1. A 2 x 3 mixed ANOVA on the surprise ratings revealed a significant main effect of Condition, $F_1(1,92) = 827.543$, $p < 0.0001$, $MS_e = .408$; $F_2(1,45) = 923.080$, $p < 0.0001$, $MS_e = .190$, whereby the surprise ratings for the confirming scenarios ($M = 1.91$, $SD = 1.29$) were rated as reliably lower than those for the disconfirming scenarios ($M = 4.69$, $SD = 1.6$). There was also a significant main effect of Group, $F_1(2,92) = 7.582$, $p = 0.001$, $MS_e = .788$; $F_2(2,45) = 5.587$, $p = 0.007$, $MS_e = .608$. Post-hoc analysis using Bonferroni adjustments showed that the Conjunction group ($M = 2.88$, $SD = 1.89$) gave significantly lower surprise ratings on average than both the Control ($M = 3.52$, $SD = 2.07$) and the Generative groups ($M = 3.35$, $SD = 1.99$, $ps < 0.01$). The interaction between Group and Condition was also significant, $F_1(2,92) = 7.441$, $p = 0.001$, $MS_e = .408$; $F_2(2,45) = 9.664$, $p < 0.0001$, $MS_e = .190$, illustrating that the Group effect was greater in the disconfirming scenarios than in the confirming scenarios.

These results are interesting for several reasons. Firstly, they demonstrate that participants were capable of generating convincing reasons for unexpected outcomes (as evidenced by the lower surprise for the Conjunction group). However, even when they did this, participants in the Generative group did not find the outcomes less surprising than participants in the Control group. Therefore, the lower ratings for the Conjunction group cannot be explained by participants' inability or disinclination to generate appropriate explanations for those events.

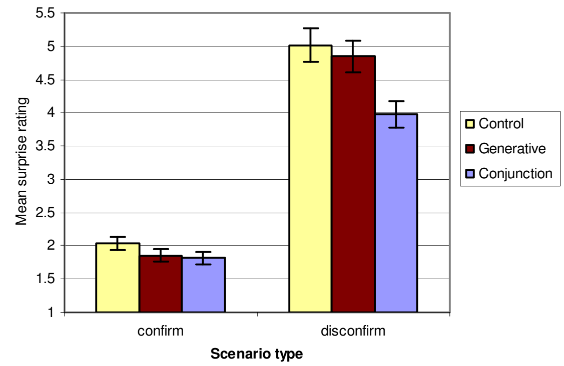

Figure 1: Surprise ratings across experimental conditions

The results leave open the possibility that, even though only one group of participants were explicitly asked to think of a reason for the unexpected events, participants in *both* the Control and the Generative group were attempting to do this. In other words, when they read the outcome *"The alarm clock failed to go off at 7am"* they tried to link this with the scenario body by means of some 'causal search' of their knowledge. This assumption is in line with most theories of discourse comprehension, which hold that people are motivated to coherently link all the events in a piece of text together using inferences (cf. Graesser et al., 1994).

Our results reveal that the reduction of surprise reported by Maguire and Keane (2006) was only manifested for the Conjunction group and not for the other two groups. At first this result seems counter-intuitive. Surely if people can think of an explanation for why something unexpected occurred, then they should not be so surprised by it? If a reader can somehow connect two events together (e.g., by hypothesising that the batteries in the alarm clock have gone flat), then they should be facilitating representational integration and thus lowering surprise. The reasons provided by the Generative group were certainly convincing, given they had the effect of lowering surprise for the Conjunction group. Consequently, the difference in surprise ratings between these groups is unusual, seeing as participants in both had access to the same information.

One explanation for this pattern of results is that hypothetical reasons are processed differently to those that are presented as part of the proposition to be evaluated. While people can easily generate a reason for an unexpected event, they may not be certain about that reason, and this may make the event seem more surprising. For any unexpected event, there will always be a number of different reasons for why that event occurred. In the alarm clock scenario for instance, an alarm clock failing to ring at 7am could be due to a power-cut, a failure of batteries, the person setting it incorrectly, or even something more bizarre like a sabotage effort. When people attempt to rationalise an event, they will be aware that multiple explanations are possible. Even if one of those explanations seems quite reasonable, it is potential that the actual sequence of events is *more* surprising than the most reasonable explanation. In other words, the less certain a given explanation is for an outcome event, the more surprising that event should be in the scenario context.

This kind of effect is evident in the way people reason about everyday events. For example, imagine you had arranged to meet someone in a city location. Ten minutes pass and they have not yet arrived – a fact that seems surprising. You imagine that perhaps they are stuck in traffic but, alternatively, you are aware that this may not be the case (e.g. they may have been involved in a car accident, which would also render them late). Eventually, when the person arrives, they explain that they *were* actually held up in traffic. Your level of surprise subsequently decreases, despite the fact that the actual reason for the delay is one that you have already considered.

It is important to note that an effect of this nature does not necessitate the generation of multiple hypotheses. On the contrary, models of hypothetical thinking maintain that people entertain only one hypothesis at a time (e.g. Evans, Over & Handley 2003). Consequently, the influence of uncertainty is likely to stem from a reduced confidence in the hypothesis rather than from the generation and consideration of alternatives. For example, in speculating why someone is late, you might feel somewhat uncertain about the *held up in traffic* explanation without having explicitly entertained any other alternatives. Thus, participants in the Generative group need not have generated more than one explanation for uncertainty to have had an effect.

General Discussion

The results of this experiment have revealed a number of interesting issues regarding the nature of surprise. Most importantly, they suggest that higher levels of uncertainty lead to higher levels of surprise. This finding is compatible with Maguire and Keane's (2006) theory of Representation-Fit which claims that surprise is based on the extent to which an event fits with a person's representation. Figure 2 illustrates this using a pair of diagrams. The first represents the Conjunction condition. Here, the explanation for the event is explicitly suggested and so surprise is based on the 'goodness of fit' between it and the existing representation. The second diagram represents the Generative condition. Here, multiple explanations are possible but since surprise is a retrospective judgement, people are aware that only one sequence of events can be correct. Given that only one of the routes linking the event to the representation can apply, the goodness of fit of each route is diluted. Although the existence of multiple explanations *increases* likelihood according to probability theory, it also serves to decrease the perceived level of fit and increase the level of surprise.

An example of this effect is a defence lawyer presenting an alibi for a defendant. Even though providing multiple possible alibis should increase the probability that the defendant is innocent, lawyers tend to present only the single strongest alibi. Because the jury is aware that only one alibi is applicable, presenting multiple explanations would actually weaken the case (Kuhn, Weinstock & Flaton, 1994). In the same way, considering multiple possible reasons for an event makes it seem more surprising than considering only the least surprising reason on its own.

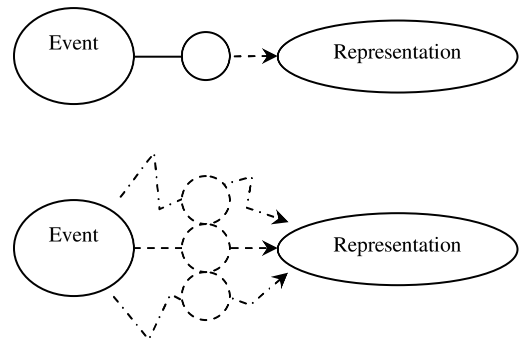

Figure 2: A graphical representation of the Conjunction (C) and the Generative (G) conditions

Surprise versus probability

Several studies have revealed a strong association between judgements of surprise and judgements of likelihood (e.g. Fisk, 2002). However, from a probability point of view, the above reasoning is clearly fallacious: a conjunction cannot be more likely than either of its constituents. For example, the conjunctive proposition of a power cut and an alarm clock failing to go off is evidently less likely than the occurrence of the latter event on its own, since this includes the possibility that the alarm clock failed to go off for other reasons. In this case, is it paradoxical that the less likely proposition should be rated as less surprising?

Given that surprise is not the same thing as probability, there is no reason why these concepts should correspond with each other. Probability takes into account the many different ways in which an event *might* occur. Thus, the more possible ways that something could happen, the more probable it will be. In contrast, surprise is a retrospective judgement concerning an event that has already occurred. Because this event can only have happened in one way, high levels of uncertainty will actually increase the potential for surprise. Accordingly, less likely events (e.g. a power cut and alarm clock failure) can actually appear less surprising than more likely events (e.g. alarm clock failure) because they minimize the potential for more surprising explanations (e.g. somebody maliciously turned off the alarm clock).

At first glance, the results of our experiments bear much resemblance to Tversky and Kahneman's (1983) study on the conjunction fallacy, raising the question of whether this might be related to surprise. In a series of experiments, Tversky and Kahneman found a pair of events was sometimes rated as more likely than the singular events on their own. For example when participants where presented with a short scenario description of a woman called Linda (who was described as an outspoken philosophy student, concerned with human rights etc), people frequently thought it was more probable that *Linda was a feminist and a bank teller* than that *Linda was a bank teller* in isolation. This occurs despite the fact that the latter option is evidently less probable than the former (i.e., since the second option covers both of the possibilities that Linda either is or is not a feminist). Discounting the possible influence of surprise, Tverksy and Kahneman (1983) explained this by virtue of a *representativeness heuristic* where people are likely to overestimate representative examples. However, recent work by Fisk (2002) has argued that this judgement fallacy is best explained in terms of Shackle's (1969) *potential surprise* theory. According to this theory, individuals' ratings of subjective probability are often influenced by the potential of that event to elicit surprise. Fisk (2002) argued that people use the potential surprise heuristic rather than Tversky and Kahneman's representativeness heuristic. In other words, people think that Linda is more likely to be a bank teller and a feminist because it is the least *surprising* possibility.

In the real world, probabilities and frequencies for particular events are rarely available. As a result, people are more likely to rely on 'gut feelings' rather than logical mathematical rules in assessing probability. It may be the case that in estimating likelihood, people often rely on judgements of how surprising an event would be were it to occur. The use of this strategy may be as a result of the singularity principle (Evans et al., 2003) which maintains that people are incapable of considering multiple hypothetical situations at the same time. Thus, the effect observed in our experiment may be of a similar nature to that observed by Tversky and Kahneman (1983). The possibility that *Linda is a bank teller and a feminist* is less surprising as an outcome than *Linda is a bank teller* because the latter includes the possibility that Linda is not a feminist. According to the singularity principle, people cannot appreciate the concept of Linda being both a feminist and not a feminist: she has to be one or the other and thus including both possibilities actually *increases* the potential for surprise. This idea can successfully explain why people tend to gravitate towards the representative sample, as reported by Tversky and Kahneman. If people base their likelihood judgements on a single hypothetical scenario then they are effectively thinking in terms of surprise rather than in terms of probability. In such cases, the most representative scenario will always be the least surprising.

Conclusion

In sum, this paper has investigated why the provision of an enabling condition decreases surprise for an unexpected event, while also decreasing probability. We found that even when participants were explicitly required to generate explanations for unexpected events, this had no significant effect on surprise ratings. Yet, when the same reasons were presented to another group, surprise was lowered. Thus, it appears that although people have the ability to infer plausible explanations for events, surprise remains high when those explanations are uncertain. In this way, a less likely event can actually be rated as less surprising than a more likely event. This effect is successfully accounted for by Maguire and Keane's (2006) Theory of Representation-Fit. Given the link between this effect and the conjunction fallacy, future study should investigate the extent to which people rely on potential surprise judgements in estimating likelihood.

Acknowledgments

This research was partially supported by the Irish Research Council for Science, Engineering and Technology.

References

Darwin, C. (1872). *The Expression of Emotion in Man and Animals.* London: J. Murray.

Evans, J.St.B.T., Over, D.E., & Handley, S.H. (2003). A theory of hypothetical thinking. In D. Hardman & L. Maachi (Eds.), *Thinking: Psychological Perspectives on Reasoning, Judgement and Fecision Making*. Chichester: Wiley.

Fisk, J.E. (2002). Judgements under uncertainty: Representativeness or potential surprise? *British Journal of Psychology, 93,* 431-449.

Graesser, A.C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*, 371-395.

Grimes-Maguire, R. & Keane, M.T. (2005a). Expecting a surprise? The effect of expectations in perceived surprise in stories. *Proceedings of the 27th Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum.

Grimes-Maguire, R. & Keane, M.T. (2005b). A cognitive model of surprise in narrative. *Kogwis: Proceedings of the Seventh Biannual Meeting of the German Cognitive Science Society.*

Kuhn, D., Weinstock, M. & Flaton, R. (1994). How well do jurors reason? Competence dimensions of individual variation in a juror reasoning task. *Psychological Science, 5,* 289-296.

Linderholm, T., Virtue, S., van den Broek, P., & Tzeng, Y. (2004). Fluctuations in the availability of information during reading: Capturing cognitive processes using the landscape model. *Discourse Processes, 37(2),* 165-186.

Maguire, R., & Keane, M.T. (2006). Surprise: Disconfirmed expectations or representation fit. *Proceedings of the 28th Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Meyer, W.U., Reisenzein, R. & Schützwohl, A. (1997). Towards a process analysis of emotions: The case of surprise. *Motivation and Emotion, 21(3),* 251-274.

Shackle, G.L.S (1969). *Decisions, Order and Time in Human Affairs.* Cambridge: Cambridge University Press.

Teigen, K.H., & Keren, G. (2003). Surprises: Low probabilities or high contrasts? *Cognition, 87*, 55-71.

Tversky, A., & Kahneman, D. (1983). Extensional vs. intuitive reasoning: The conjunction fallacy in probability judgement. *Psychological Review*, *90 (4),* 293-315.

Zwaan, R.A. & Radvansky, G.A. (1998). Situation Models in Language Comprehension and Memory. *Psychological Bulletin, 123 (2),* 162-185.