Surprise: Disconfirmed Expectations or Representation-Fit?

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

Abstract

Surprise for a particular event is often thought to correspond to the degree to which that event deviates from an expected schema. However, in this paper we present two novel experiments that challenge this view. Participants were asked to rate how surprised they would be if an event, or series of events, followed on from a number of short scenarios. In one condition these events confirmed an expected outcome, while in another they contradicted this outcome. A third condition included a potential enabling factor along with the unexpected outcome. The results show that, even when events deviate significantly from an established schema, surprise is lower if people have a means of integrating that event into their representation. This finding is consistent with our theory of Representation-Fit, which asserts that a person's level of surprise for a given event can be determined by how well that event is supported by the prior discourse.

Keywords: Surprise, expectation, representation, discourse.

Introduction

Imagine how you would feel if you walked into your house to discover a group of strangers huddled around your dining room table playing cards and smoking cigars. Alternatively, consider how you might react if you booked your holiday to Outer Mongolia, only to bump into your next-door neighbour strolling along a deserted mountain path. While such events are unlikely to happen to us in real life, we have no difficulty in imagining how we would feel if they were to occur: essentially, we would be surprised. Surprise is classed as one of the most basic and universal of human emotions (Darwin, 1879; Ekman, 1979), and as well as being associated with a distinct subjective and physiological response, it is also known to have some important cognitive manifestations (e.g., Fisk, 2002; Meyer, Reisenzein & Schützwohl, 1997; Teigen & Keren, 2003). 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 (Schützwohl and Reisenzein, 1999). The purpose of such a reaction is to discover *why* the surprising event transpired, so that a similar event might be anticipated in future circumstances.

While the subjective phenomenon of surprise is well established, deciding what counts as a surprising *event* is a more challenging task. Our aim in this article is to shed some light on this issue, and more specifically, to assess the claim that surprise follows disconfirmed expectations. We also put forward an alternative hypothesis: namely, that surprise for a given event is based on how well that event can be integrated into a person's discourse representation.

Probabilities, expectancy and schema-discrepancy

The most intuitive way of describing a surprise 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 (Shackle, 1969; Teigen & Keren, 2003). Consider a lottery draw, where six numbers are selected at random. It is unlikely that you would be surprised if the seemingly arbitrary selection {3, 7, 15, 18, 24, 37} was drawn, despite the extremely low probability of this particular sequence occurring. In contrast, if the numbers selected were {1, 2, 3, 4, 5, 6}, you would undoubtedly be far more surprised, even though this sequence has exactly the same probability of occurring as the former. In view of this, Shackle (1969) proposed that, rather than defining surprise in terms of *actual* probability, it can be more accurately described as a reflection of *subjective* probability. Hence in the above example, the latter combination of numbers has a much lower subjective probability of occurring, due to its overtly meaningful pattern. On the other hand, we know that some 'random' combination of numbers is far more typical, making the first sequence appear less surprising. This is in line with Tversky and Kahneman's (1983) finding that people will often make systematic errors in probability judgements if they perceive certain events as being more 'representative' than others.

Nevertheless, while a number of studies have found a strong correlation between subjective probability and surprise (e.g. Fisk, 2002), this relationship is not always quite so straightforward. One important difference is that judgements of probability are usually given *before* the event in question, whereas surprise is experienced *after* that event has taken place. Studies on hindsight bias have illustrated that people will often adjust how likely they thought an event was after that event has occurred (Hawkins & Hastie, 1990). It follows that posing the question *"How likely is it to rain today?"* might elicit a different response to *"How surprised are you that it rained today?"* In light of this, Teigen and Keren (2003) suggest that surprise at a given event might be more accurately explained in terms of its subsequent comparison with an alternative event. Specifically, their *Contrast Hypothesis* maintains that surprise is dependent on the degree of disparity between an observed and an expected outcome. Investigating this hypothesis, Teigen and Keren (2003) 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 one 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 a chance at winning the race, no expectation is contradicted if Erik wins. A similar theory was put forward by Meyer et al (1997): their *Cognitive-Psychoevolutionary* model proposes that surprise is experienced when expectations based on existing schemas fail to be confirmed. Surprise is therefore considered as a reaction to a schema-discrepant event.

Henceforth, we will refer to the above theories as encapsulating the *Expectation-Disconfirmation* hypothesis. At first blush, this hypothesis seems intuitive. Consider, for example how surprised you would be if you threw a ball up in the air and it failed to fall to the ground. This is an extreme case of an expectation (i.e., that the ball will drop due to gravity) being contradicted. However, it may not *always* be the case that disconfirmed expectations like this lead to such a high level of surprise. Other factors in event comprehension, in particular the adaptability of one's existing representation, may mitigate the extent of this phenomenon: in short, if a person can account for *why* an expectation was disconfirmed, then they might not be so surprised by it. This is the central premise of our theory of Representation-Fit, which is outlined in more detail below.

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. 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 conceived 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 leave your house to discover that your car is no longer in the driveway. This is obviously an unexpected, or schema-discrepant, event. However, if you remember that you left the car in to be serviced yesterday morning, you will be not be so surprised, since a satisfactory explanation for this unexpected event has been identified. A study conducted by Grimes-Maguire and Keane (2005a) offered substantive empirical evidence for this perspective. We 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 was supported by the prior discourse. We 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.

However, while the findings of the above study are in accordance with the theory of Representation-Fit, they do not necessarily oppose the Expectation-Disconfirmation view because no expected outcomes were explicitly disconfirmed in the scenarios examined. Consequently, the experiments detailed in the current paper investigate the influence of representation-fit in cases where the discourse clearly contains a schema-discrepant event.

Overview of experiments

In the following experiment, we devised a series of event sequences that instantiated an expectation for a particular outcome. This outcome was subsequently confirmed or disconfirmed in two separate conditions. In addition, a third condition was included where the disconfirmed outcome was presented alongside an enabling factor for that outcome. The Expectation-Disconfirmation hypothesis does not predict that surprise ratings will be affected by an enabling factor, as these events still contradict the expected sequence of events. However, in accordance with the theory of Representation-Fit, we hypothesised that participants would be less surprised when provided with the enabling factor for the disconfirming outcome, as this event should facilitate representation integration. Experiment 2 examines another potential explanation for the observed findings.

Experiment 1

In order to explore the Expectation-Disconfirmation hypothesis of surprise, a number of event sequences were designed that strongly suggested a particular outcome (see Table 1 for a sample material). For each of these scenarios, four hypothetical outcomes were generated, corresponding to the following four conditions:

1. *Expectation confirmed* (C) described an event that corroborated the expected outcome (e.g. in Table 1, the alarm clock ringing at 7am supports the expectation that is implied by the scenario body).
2. *Expectation disconfirmed* (D) described an event that contradicted the expected outcome (e.g., the alarm clock failing to ring directly opposes the intuitive expectation).
3. *Expectation disconfirmed and enabling event* (D+enable) described the same disconfirming event as in the previous condition alongside a potential enabling factor for that event (e.g. a power-cut during the night explains why the alarm clock might fail to ring).
4. *Control* (D+control) included the same disconfirming event as in the latter two conditions in accompaniment with an irrelevant event (e.g., a good night's sleep should not influence the functioning of an alarm clock). This condition was intended to control for the enabling event in the D+enable condition.

According to the Expectation-Disconfirmation hypothesis, while the C condition should be judged as unsurprising, there should be no difference in the surprise ratings given for the latter three conditions. This is because all three contradict the expected outcome by means of the same disconfirming event. In contrast, we hypothesised that surprise ratings would follow the trend C < D+enable < D. In line with the Representation-Fit perspective, the D+enable condition should be perceived as less surprising than the D condition, since it offers the reader a rationalisation for why the unexpected event might have occurred. This condition should therefore be more easily integrated into the discourse representation. We did not expect surprise ratings from the D condition to differ from the D+control one.

Method

Participants Fifty-two undergraduate psychology and maths students (all native English speakers) from University College Dublin voluntarily participated in this experiment.

Materials Sixteen event sequences describing simple everyday scenarios were generated, each of which instantiated an expectation that was either confirmed or disconfirmed in the four separate conditions as outlined above (see Table 1 for an example).

Design The conditions were counterbalanced using four different groupings of scenarios and conditions. Each participant was presented with one of these groupings which contained the 16 scenarios. Participants thus read a total of four scenarios from each condition.

Procedure Participants were given a booklet containing the 16 scenarios, where instructions for the experiment were displayed on the first page. Participants were informed that they would be presented with a number of everyday scenarios and that they should rate how surprised they would be *if* a certain event, or series of events, followed on from that scenario. An example was then provided to demonstrate the nature of the task. Each scenario body was presented on a separate page, followed by the words *"How surprised would you be if: X"*, where X referred to the event, or series of events, corresponding to one of the four experimental conditions. Participants were required to rate their level of surprise for this event on a 7-point scale, with 1 indicating that they would be very unsurprised if the event occurred, and 7 indicating that they would be highly surprised if the event occurred. The scenarios were presented in a different random order to each participant.

Table 1: Scenario from Experiment 1. All participants read the scenario body and rated surprise for one of the four conditions. Mean surprise ratings are also displayed.

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….?	***Surprise***
1	The alarm clock woke her up at 7am ***(C)***	1.88
2	The alarm clock failed to ring at 7am ***(D)***	4.95
3	There was a power-cut during the night and the alarm clock failed to ring at 7am ***(D+enable)***	4.46
4	She had a quiet, good night's sleep and the alarm clock failed to ring at 7am ***(D+control)***	5.04

Results & Discussion

The results support the Representation-Fit account and argue against the claim that disconfirmed expectations will consistently lead to an equally high level of surprise. When participants are provided with a reason for why an unexpected event might occur, their surprise is lower than when they are simply presented with this event on its own.

The mean surprise ratings for each of the four conditions are displayed in Table 1. A one-way ANOVA, repeated measures revealed a significant difference in these ratings across conditions, both by participants, $F_1(3,153) = 134.401$, $p < 0.0001$, $MS_e = .858$, and by materials, $F_2(3,45) = 107.801$, $p < 0.0001$, $MS_e = .331$. Post-hoc analysis using Bonferroni adjustments (all ps < 0.0083) showed that surprise ratings for the C condition ($M = 1.88$, $SD = 1.38$) were significantly lower than the ratings in the other three conditions. This finding is intuitive, as outcomes that are consistent with expectations will inevitably be judged as less surprising than those that run contrary to expectations. The surprise ratings for D ($M = 4.95$, $SD = 1.77$) and D+control ($M = 5.04$, $SD = 1.74$) were not significantly different, which demonstrates that the addition of an irrelevant event had no effect on perceived surprise. However, ratings for the D+enable condition ($M = 4.46$, $SD = 1.72$) were significantly lower than those for both the D and D+control conditions. This finding undermines the Expectation-Disconfirmation hypothesis, as it suggests that the same unexpected event will not always be judged as equally surprising. Participants provided with a rationale for *why* this unexpected event might occur were less surprised than participants presented with that unexpected event in isolation. This suggests that subjective surprise is based on the ease of representational integration rather than on factors such as probability or schema-discrepancy.

In addition to the ANOVA we performed a trend analysis to ascertain if our predicted trend of C < D+enable < D held true. A Page's L corroborated this trend across both participants, $L(2) = 703.5$, $p < 0.0001$, and materials, $L(2) = 220$, $p < 0.0001$. This confirms that participants were more likely to rate events in the D+enable condition as less surprising than those in the D condition. In a further investigation, we divided participants' surprise ratings into categories of low (1-2), medium (3-5) and high (6-7). We then performed a chi-square analysis to see if there was a relationship between surprise category and condition. The association between these measures was significant, $\chi^2(6, N = 832) = 370.931$, $p < 0.0001$. Examination of the standardised residuals (set at +/- 2) revealed that participants were most likely to give a low surprise rating for C, a mid surprise rating for D+enable and a high surprise rating for D and D+control.

In summary, this experiment has shown that surprise for an unexpected event is mitigated when a means for rationalising that surprise is made available. Although this effect was not great enough to lower surprise ratings to the base level of an expectation-confirming event, we have demonstrated that providing participants with a reason for a schema incongruent event can significantly attenuate their level of surprise.

Experiment 2

It is possible that the above results do not contradict the Expectation-Disconfirmation hypothesis. Another potential interpretation of the findings could be that participants reading scenarios in the D+enable condition were somehow altering their expectation for the outcome *after* reading the enabling event. This would explain why surprise ratings for the unexpected event were lower in this condition than they were when the disconfirming event was presented in isolation (as in D). For instance, participants reading the scenario in Table 1 should have expected the alarm clock to go off at the appointed time. However, when judging the possibility that *"there was a power-cut during the night and the alarm clock failed to ring at 7am"*, they may have based their surprise ratings for the unexpected outcome ("*the alarm clock failed to ring"*) in the context of the enabling event ("*there was a power cut during the night"*). In other words, participants might have viewed this enabling factor as a given and only rated surprise for the unexpected event in light of this. If this were true, the intended expectation would have been invalidated in the D+enable condition (i.e. given a power cut, a radio alarm clock is not likely to function).

In the following experiment, we investigated this possibility with the aim of refuting the argument. If participants in Experiment 1 were really interpreting enabling events as part of the scenario and thereby altering their expectation, then presenting the enabling event within the scenario body should make no difference to perceived surprise. In opposition to this view, we assume that this manipulation should considerably alter surprise ratings for the unexpected outcome. That is, if the enabling event is portrayed within the scenario body, then it will be integrated into the reader's representation from the onset. When asked to rate surprise for the 'disconfirming' event, ratings should be lower since supportive information for this event will be evident in the representation (i.e., when readers *know* that there has been a power-cut, they should not be surprised by the possibility that the alarm clock will fail to ring). In contrast, when the enabling event is presented as part of the hypothetical event sequence (as in the D+enable condition), surprise should be greater because that sequence will not be supported by the existing representation (i.e., the occurrence of a power-cut during the night and the alarm clock failing to ring still go against the expected sequence of events).

Table 2: Sample scenario from Experiment 2. Participants all read scenario body and additional sentence where appropriate (these are shown in brackets within the scenario body).

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. [There was a power-cut during the night.*] [She had a quiet, good night's sleep.**] *only presented in *D(enable known)* condition ** only presented in *D(control known)* condition
How surprised would you be if….?
The alarm clock failed to ring at 7am *D, D(enable known), D(control known)*
There was a power-cut during the night and the alarm clock failed to ring at 7am *D+enable*

In this experiment, we compared surprise ratings for the original D and D+enable conditions, along with a novel condition, D(enable known), in which the enabling event was presented within the scenario body. An appropriate control condition was also included which contained an irrelevant event in the scenario body, D(control known). For both of these conditions we wished to see how surprising participants would find the 'disconfirming' event on its own. This contrasts with the D+enable condition, where participants must rate surprise for the enabling event and the disconfirming event together. Table 2 displays a sample scenario elucidating these various conditions.

The alternative account of the results from Experiment 1 predicts the following ordering for the three disconfirming conditions: D+enable = D(enable known) < D. Conversely, the Representation-Fit account predicts that surprise ratings will differ across all three conditions, with the condition where the enabling factor is in the scenario body being the least surprising: D(enable known) < D+enable < D.

Method

Participants Fifty-seven undergraduate computer science students from University College Dublin voluntarily took part in this experiment. All were native English speakers. Data from two of these participants were discarded prior to analysis due to a failure to complete the experiment.

Materials, Procedure & Design. The 16 event sequences from Experiment 1 were modified for the purposes of this experiment. Four versions of the scenarios were generated according to the four conditions as outlined above (see sample scenario in Table 2). The procedure and design were identical to those in Experiment 1. Each participant thus rated four scenarios from each condition.

Results & Discussion

The mean surprise ratings for the four conditions are displayed in Figure 1. These results are in line with our predictions, in that events in the D(enable known) condition were judged as the least surprising (M = 3.58, SD = 1.85), followed by those in the D+enable condition (M = 4.19, SD = 1.62), and lastly by those in the D (M = 4.91, SD = 1.68) and D(control known) conditions (M = 4.99, SD = 1.68). A one-way ANOVA, repeated measures revealed a significant difference in surprise ratings across these conditions, both by participants, $F_1(3,162) = 38.125$, $p < 0.0001$, $MS_e = .638$, and by materials, $F_2(3,45) = 19.802$, $p < 0.0001$, $MS_e = .364$. Post-hoc analysis using Bonferroni adjustments (all ps < 0.0083) showed that surprise ratings for the D(enable known) condition were significantly lower than the ratings in the other three conditions. As predicted by the Representation-Fit theory, these results clearly demonstrate that when the enabling event forms part of the existing representation, surprise is lower. The results also undermine the possibility that participants' expectations were influenced by the enabling event in the D+enable condition, since items in this condition were rated as reliably more surprising than those in the D(enable known) condition. However, ratings for the D+enable condition were significantly lower than those in the D and the D(control known) conditions which replicates the findings from Experiment 1. This reinforces the claim that the inclusion of an enabling factor decreases surprise for a disconfirming event. The D(control known) condition was not reliably different to its matching D condition.

A Page's L confirmed the trend in surprise ratings predicted by the Representation-Fit account of D(enable known) < D+enable < D, across participants, $L(2) = 728.5$, p < 0.0001, and materials, $L(2) = 216$, p < 0.0001. A chi-square analysis examining the frequencies of the surprise rating categories (low, medium and high) within the four conditions was also significant, $\chi^2(6, N = 873) = 105.038$, $p < 0.0001$. Standardised residuals revealed that participants were most likely to give a low surprise rating for scenarios in the D(enable known) condition, a mid surprise rating for those in the D+enable condition and a high surprise rating for the D and D(control known) conditions.

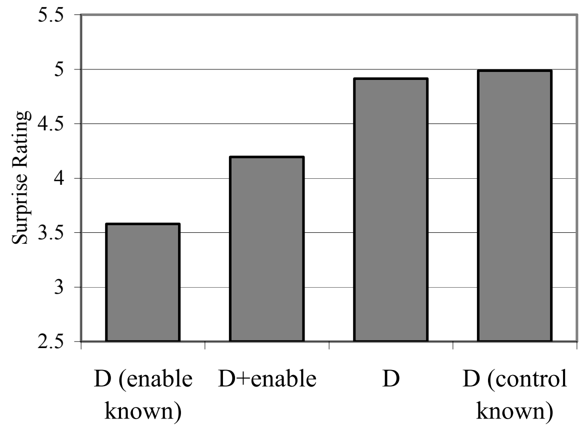

Figure 1: Mean surprise ratings for the conditions in Exp 2

In sum, the results of this experiment are consistent with those of Experiment 1 and vindicate our original experimental paradigm by showing that the availability of explanatory factors for a surprising event can greatly facilitate representational integration. This again indicates that the assessment of surprise involves more than mere expectation-disconfirmation or schema discrepancy.

General Discussion

In this paper we have investigated the claim that disconfirmed expectations, or schema-discrepant events, lead to a consistently high level of surprise (e.g. Meyer et al, 1997; Teigen & Keren, 2003). While this is an appealing theory, results from two novel experiments suggest that the ease of integration, or what we have termed Representation-Fit, may offer a more accurate perspective from which to consider this phenomenon. These experiments showed that people will judge an unexpected event as less surprising when a potential enabling factor for that event is made available. This effect can be observed even when events deviate significantly from a dominant schema by conflicting strongly with an expected sequence of events. In this section, we will discuss the implications of these results for our understanding of surprise, before outlining a cognitive model designed to explain the phenomenon.

An important motivation for these experiments was to see whether the experience of surprise could be related to that of expectation. These two concepts are frequently used interchangeably but they may not be analogous (see Grimes-Maguire & Keane, 2005a, 2005b). While our theory of Representation-Fit does not see surprise as being dependent on a person's level of expectation, this variable can obviously play a role in surprise judgements. Evidently for example, if a certain outcome is expected, that event will be effortlessly integrated into a person's representation and will consequently be perceived as very unsurprising. However, as we have shown here, if an event goes *against* an expectation, surprise can be lowered if some rationalisation for that event is either provided or detected. By this token, an unexpected event should no longer seem so surprising when the reasons for it are understood. It is not clear whether theories such as that proposed by Meyer et al (1997) can adequately explain this effect.

At first glance, the results of our experiments bear much resemblance to Tversky and Kahneman's (1983) notion of *representativeness*, which is the idea that two events will appear more likely than one event on its own if those two events are perceived to be more representative of some norm. However, we have already mentioned that surprise cannot simply be defined as a function of such subjective probabilities since it is a *retrospective judgement* rather than an estimation of likelihood. Teigen and Keren (2003) also demonstrated that these two variables often yield differential judgements by people. In light of this, while Tversky and Kahneman (1983) found that event sequences appear more probable when they converge on the norm, we claim that event sequences seem less surprising when people are presented with a means of *adjusting* their norm.

Cognitive model of surprise judgements

Maguire, Costello and Keane (2006) have devised a model of surprise judgements that is in line with the theory of Representation-Fit. Sharing some similarities with a model of plausibility judgements proposed by Connell and Keane (2006), the computational instantiation of our theory outputs a surprise rating for the final event in a number of short scenarios by means of two distinct phases. Firstly, the *Integration phase* entails building a coherent representation of the scenario. During this phase, the objective is to integrate any incoming events into the discourse representation in the easiest way possible. The *Analysis phase* follows when the model assesses level of surprise for a particular event. This is done by ascertaining the degree of representation-fit between the outcome and the events in the prior discourse representation. Recent simulations have revealed a strong correspondence between human data and outputted surprise ratings. One important assumption of the model has been supported by the current experiments – that people actively search for consistent information or enabling factors when assessing surprise for a certain event.

Conclusion

Despite being a subject of investigation for some time, there is still much to be discovered about the cognitive basis of surprise. In this article, we have demonstrated that subjective surprise is not a straightforward function of the degree of contrast between an observed and an expected outcome. We have illustrated that it is not the extent of schema discrepancy which determines the level of surprise, but rather the ease with which the discrepant event can be rationalised and subsequently integrated into the discourse representation. Existing theories of surprise may need to be refined to acknowledge this representational influence.

Acknowledgments

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

References

Connell, L. & Keane, M.T. (2006). A model of plausibility. *Cognitive Science, 30(1),* 95-120.

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

Ekman, P. (1972). *Emotion in the Human Face: Guidelines for research and an integration of findings.* New York: Lawrence Erlbaum.

Fisk, J.E. (2002). Judgments 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.*

Hawkins, S. A. & Hastie, R. (1990). Hindsight: Biased judgements of past events after the outcome events are known. *Psychological Bulletin, 107,* 311-327.

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., Costello, F. & Keane, M. T. (2006). A cognitive model of surprise judgements. *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.

Schützwohl, A. & Reisenzein, R. (1999). Children's and adults' reactions to a schema-discrepant event: A developmental analysis of surprise. *International Journal of Behavioral Development, 23,* 37-62.

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 judgment. *Psychological Review*, 293-315.

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