Available online at www.sciencedirect.com
SCIENCE DIRECT®
Cognition 87 (2003) 55–71
COGNITION
www.elsevier.com/locate/cognit

Surprises: low probabilities or high contrasts?

Karl Halvor Teigen[a,*], Gideon Keren[b]

[a]*Department of Psychology, University of Oslo, P.O.B. 1094 Blindern, N-0317 Oslo, Norway*
[b]*Faculty of Technology Management, Eindhoven University of Technology, P.O.B. 513, 5600 MB Eindhoven, The Netherlands*

Received 1 June 2001; received in revised form 27 February 2002; accepted 29 October 2002

Abstract

Outcome expectations can be expressed prospectively in terms of probability estimates, and retrospectively in terms of surprise. Surprise ratings and probability estimates differ, however, in some important ways. Surprises are generally created by low-probability outcomes, yet, as shown by several experiments, not all low-probability outcomes are equally surprising. To account for surprise, we propose a *contrast hypothesis* according to which the level of surprise associated with an outcome is mainly determined by the extent to which it contrasts with the default, expected alternative. Three ways by which contrasts can be established are explored: contrasts due to relative probabilities, where the obtained outcome is less likely than a default alternative; contrasts formed by novelty and change, where a contrast exists between the obtained outcome and the individual's previous experience; and contrasts due to the perceptual or conceptual distance between the expected and the obtained. In all these cases, greater contrast was accompanied by higher ratings of surprise.

Keywords: Surprise; Subjective probabilities; Expectations; Contrast

1. Introduction

In theories of emotion, surprise is often described as one of the six or eight most basic emotions (Ekman & Friesen, 1971; Plutchik, 1991), particularly because of its unique pattern of facial expressions (Young et al., 1997). At the same time it differs from most other emotions by being instigated both by positive and negative events, the most salient characteristic being their unexpectedness, rather than their valence (Roseman, Antoniou, & Jose, 1996). Surprise has accordingly been described as a cognitive affect (Buck, 1999). In studies of infant cognition, expressions of surprise are often used to indicate beliefs

* Corresponding author. Tel.: +47-22-845187; fax: +47-22-845175.
E-mail address: k.h.teigen@psykologi.uio.no (K.H. Teigen).

0010-0277/02/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0010-0277(02)00201-9

about objects (Baillargeon & Hanko-Summers, 1990; Mareschal, Plunkett, & Harris, 1999).

If surprise is, by definition, elicited by unexpected events, it may appear unnecessary to investigate this relation empirically (Smedslund, 2002). Following this line, some investigators claim they are studying surprise when they in fact study expectancy confirmation/disconfirmation (Petty, Fleming, Priester, & Feinstein, 2001), or responses to events outside the anticipated range of outcomes (Coughlan & Connolly, 2001). Expectations, however, differ in degree and the way they are conceptualized. In cognitive psychology, expectations have often been quantified in terms of probabilities. According to this interpretation, unexpected events can simply be described as low-probability outcomes. This suggests a straightforward negative relationship between probability and surprise, constituting a link between a central concept in the cognitive psychology of judgment and decision making (cf. Kahneman, Slovic, & Tversky, 1982) and one of the basic emotions.

The link between probabilities and surprise has not received much attention in the research literature, perhaps because of its assumed obviousness. Occasionally, the term surprise is used synonymously with low probability (Mellers, Schwartz, & Ritov, 1999; Slovic & Fischhoff, 1977). In his seminal book, "Applications of Information Theory to Psychology", Fred Attneave (1959) defined the *surprisal* of an event in terms of its informational value, as a simple function of the probability of an event. However, he warned the reader that "events with equal surprisal values may not be equally surprising to an observer". A complementary perspective was adopted by Shackle (1969), who suggested that the subjective probability of an event is actually mediated by its potential surprise value (i.e. the extent to which the outcome includes low-probability components). Implications of this model of probability judgments were recently explored by Fisk and Pidgeon (1997, 1998). None of the above-mentioned authors, however, appears to have performed empirical studies in which participants were explicitly asked to assess surprise.

Probabilities and surprise differ not merely on the cognitive/emotional dimension, but also by being assessed at different points in time. Probabilities are most meaningfully evaluated before the outcome, whereas surprise is typically experienced after the outcome is known. Notwithstanding, it is possible to ask people to anticipate how surprised they think they will be by a particular outcome. Conversely, in hindsight studies people are asked to form probability judgments in retrospect. In a cross-cultural study of surprise, Choi and Nisbett (2000) found comparable ratings of probabilities and anticipated surprise. They also found a similar pattern of results for rated surprise and probabilities in hindsight. Even more relevant for our present concern is a recent study by Reisenzein (2000) who asked participants to rate their confidence in their preferred answers to a 35-items multiple-choice test, and then to rate their experienced surprise after the true answers were disclosed. Confidence ratings for individual subjects correlated on the average -0.78 with their rated degree of surprise, showing a very close association between low probability and degree of surprise, or, in Reisenzein's terminology, between the cognitive and the experiential component of the surprise syndrome.

There are thus good reasons to believe that, *ceteris paribus*, high surprise ratings are associated with low probabilities, and vice versa. Still, the surprise may not be due to its low probability per se, but its degree of contrast with the more likely or 'normal' disconfirmed expectation. This interpretation is in line with the observation by Miller, Turnbull,

and McFarland (1990) that two equally improbable events may not be equally normal, and hence not equally surprising or suspicious. In their example, drawing a chocolate cookie from a jar with one chocolate and 19 oatmeal cookies appears more striking (and hence more suspicious) than drawing it from a jar with ten chocolate and 190 oatmeal cookies. The first event is more non-normal and surprising because it can happen in only one way, whereas the second can be imagined to happen in several different ways.

In the present article we ask whether outcomes with identical probabilities will always lead to the same degree of surprise, or whether, sometimes, unlikely outcomes are less surprising than more likely ones. By answering this question we hope to obtain some further insight into the nature of surprise and the variables that may affect it.

In the following, a series of experiments is presented in which participants rate the degree of surprise associated with outcomes with known or estimated probabilities. From the results, we are led to conclude that rated surprise and subjective probabilities are not identical. Instead, we offer a contrast interpretation of surprise according to which outcomes are surprising insofar as they conflict with other, dominant expectations. Contrasts between expectations and actual outcomes can come forth in more than one way. In the present article, we analyze and test empirically three different forms by which a contrast and the ensuing surprise can emerge.

Surprises are generally created by low-probability outcomes, yet not all low-probability outcomes are necessarily surprising. For surprise to take place, alternative outcomes with high expectations (that stand in contrast to the low-probability outcome) should endure. Thus, an astrophysicist may be extremely "surprised" to learn that the moon is one billion years older than previously thought, whereas a layman may declare himself not surprised at all; not because the new estimate was expected, but because it did not conflict with any previously held beliefs. Similarly, most people would experience little surprise when ticket No. 14 237 is drawn in a lottery with 100 000 tickets sold, despite its exceedingly low probability, as long as no other number is more strongly expected. The most surprised player will probably be the winner, who is more inclined to partition the outcome space in 'winning' and 'not winning', and can compare the small odds of winning to the much larger odds of the alternative.

Contrasts can also be created by dominant expectations arising from accumulated personal experiences, which have established a standard of comparison for future outcomes. Novel or unfamiliar outcomes will then appear more surprising than more familiar ones, even in cases where they are not less likely. For instance, a young researcher may become more surprised both by an acceptance and a rejection letter from a journal editor, compared to an experienced researcher, whom "nothing can surprise".

Finally, degree of surprise will not merely reflect the improbability of the actual outcome, but also by how different or dissimilar it is from the expected alternative. The more the actual outcome deviates from a prototypical expected outcome, the larger the contrast and the ensuing surprise.

The contrast interpretation of surprise is congruent with the Meyer, Niepel, Rudolph, and Schuetzwohl (1992) definition of surprise as the reaction to a "schema-discrepant event". Previous authors, however, seem not to have made an explicit distinction between (a) the fact that a particular outcome is unexpected (i.e. has a low probability of occurrence), and (b) that it contrasts with, and hence disconfirms, a dominant expectancy.

"Unexpectedness" appears to be an ambiguous term, because it refers to the characteristics of both the actual outcome and to its competing non-occurring alternative(s).

The present studies test the three different implications of the contrast interpretation of surprise. In Section 2, we investigate how the surprise of a low-probability outcome is affected by the probability of the other alternatives. We conjecture that a low-probability outcome will be perceived as more surprising when contrasted with a dominant, high-probability alternative, than when the probabilities of all outcomes are uniformly low. Accordingly, the surprise associated with an outcome is determined by the relative, rather than the absolute probabilities involved.

In Section 3, two studies are reported showing that an outcome is regarded as surprising to the extent that it conflicts with an individual's previous experience in the domain under consideration. Even events that are estimated (rightly or wrongly) as likely can be potentially surprising if, at the same time, they appear novel and unfamiliar. Consequently, some "probable" but novel outcomes will be judged as more surprising than other, less probable, but more familiar outcomes.

Finally, a target outcome can appear surprising not just by being less probable and less familiar than the expected outcome, but also by being dissimilar. The contrast hypothesis predicts that, for instance, a card player who believes that nobody can beat his nine of hearts will be more surprised by an unexpected king than by an equally unexpected ten of hearts, simply because the king is more strikingly different from the kind of cards he expects to see on the table. This conjecture is tested by two experiments in Section 4 of the article.

2. Surprise as determined by the relative probabilities of alternative outcomes

The surprise of an unexpected target outcome T can be construed in two different ways: (1) T is surprising because it was not expected, or (2) T is surprising because *something else* was expected more strongly. According to a contrast interpretation of surprise, the second formulation is more natural and more potent to yield surprise than the first one. Both formulations will, however, lead to indistinguishable predictions with dichotomous, complementary outcomes. If one is surprised to learn that A won the match against an assumed stronger opponent B, it is difficult to tell whether the surprise is due to the poor odds of A or the enhanced odds of B. Given several alternative outcomes (e.g. several competitors), these two interpretations can be separated and put to test.

The importance of relative probabilities has previously been demonstrated in the realm of verbal probabilities and preferences among gambles as an "Alternative-outcomes effect" (Windschitl & Wells, 1998). Windschitl and Young (2001) have proposed that people evaluate their winning chances according to a comparison heuristic, in which special attention is being paid to their chances relative to other strong players (rather than to the aggregate of other players). In a similar vein, Teigen (1988, 2001) found that people are willing to describe the chances of a target alternative in the context of several equiprobable alternatives with high-probability positive phrases, like 'good chance', 'entirely possible', and 'not improbable', rather than with low-probability negative phrases, like 'rather improbable', 'quite uncertain', and 'somewhat doubtful' (the

equiprobability effect). Hence, in a multiple-alternative context, numeric low-probability outcomes can be described as rather plausible. The present studies extend this issue to ratings of surprise. We hypothesize that a low-probability outcome will be perceived as less surprising in a context of multiple low-probability alternatives, than in comparison to one other dominant event.

Experiment 1 was designed to show that an outcome of a die game with a known probability of 1/3 is perceived to be more surprising in the context of one, more probable alternative, than in the context of several, less probable outcomes. In Experiment 2, we collect surprise ratings and probability estimates of the outcome of an athletic contest, in which the protagonist beats either one favorite, or several equal opponents. Both experiments show that surprise is dependent on the probability of the target outcome, relative to the alternative outcomes.

2.1. Experiment 1

2.1.1. Method

2.1.1.1. Participants The participants were 153 students from the University of Utrecht recruited by an advertisement and paid 12.50 Dutch guilders (approximately $7.00) for completing these and other decision-making tasks.

2.1.1.2. Design and procedure All participants were presented with a scenario describing a die game with monetary payoffs. Participants in the *single-alternative* condition read the following description:

> Tom and Peter are playing dice for money. At the end of the day, Tom has 50 guilders left and Peter has 100 guilders. They agree to let one final throw of the die decide who will take all the money home. If a 5 or a 6 is thrown, Tom will be the winner. If 1, 2, 3, or 4 is thrown, Peter will win it all.

In the *multiple-alternatives* condition, Tom was playing with four other friends. Tom has 50 guilders left and the others have 25 guilders each. Tom is said to win the entire pot if the die shows 5 or 6. If the die shows a 1, 2, 3, or 4, Peter, Jimmy, Willem, and Robert, respectively, will win the pot (each with a probability of 1/6). Thus, Tom's chances were the same in both conditions, but the number of alternatives (and their probabilities) were different.

In both conditions participants were asked to estimate Tom's winning chances. Subsequently, they were told that the die was thrown, and Tom turns out to be the winner. Half of the participants in each condition were asked to indicate (on a 1–7 scale) how surprised they would be if they were Tom, and the other half were instructed to assess how surprised they would be if they were Peter (who did not win). Thus, the experiment consisted of a 2 (single vs. multiple alternatives) × 2 (Tom's vs. Peter's surprise) between-subjects design.

2.1.2. Results

As shown in Table 1, surprise ratings for the single-alternative scenario were, on the average, much higher than the ratings for the multiple-alternatives scenario. A logit

Table 1
Mean surprise ratings (1–7) of Tom winning a die game with $P = 1/3$ (Experiment 1)

	Single-alternative condition (n)	Multiple-alternatives condition (n)
Tom's (winner's) surprise	5.03 (36)	3.98 (49)
Peter's (loser's) surprise	4.44 (34)	3.23 (34)
Mean	4.73	3.60

ANOVA indicates that the difference between the two conditions was highly significant ($z = 5.41$, $P < 0.001$). In addition, the mean surprise ratings for Tom were higher than the comparable ratings for Peter, a difference that was also statistically significant ($z = 3.31$, $P < 0.001$). The interaction between these two effects was not significant.

The results indicate that the surprise of an outcome with $P = 0.33$ (Tom's probability to win) will depend heavily upon the probability of the other alternatives, and also upon its valence, being more surprising for the winner than for the loser. These results also confirm our previous finding (Teigen & Keren, 2002) that in a lottery context, successes are more surprising than failures, as Tom, the winner, is believed to be more surprised than Peter, the loser.

2.2. Experiment 2

2.2.1. Method

2.2.1.1. Participants The participants were 178 students from the University of Utrecht recruited in the same way as participants in Experiment 1.

2.2.1.2. Design and procedure All participants were presented with a scenario (adapted from Teigen, 2001) describing Erik, who is competing in a 5000 m track race against ten other athletes. Subjects in the *single-alternatives* condition received the following information:

> It is a tough race. When there are only two more laps (800 m) to go, Erik is on the second place, lagging behind the front runner by 20 m, with the rest of the runners far behind. For quite a while, it appears that the distance between Erik and the athlete in the leading position has remained unchanged.

For subjects in the *multiple-alternatives* condition the situation was described as follows:

> It is a tough race. When there are only two more laps (800 m) to go, all athletes are still forming one group.

Participants in both conditions were asked (1) to assess the probability that Erik will win, and (2) to rate how surprised they would be if Erik won the race. Half the participants in each of the conditions assessed the probability before rating surprise, and the other half

Table 2
Mean probability estimates and surprise ratings in the track race scenario in Experiment 2

Order of ratings	Single-alternative condition (Erik vs. one better runner)		Multiple-alternatives condition (Erik vs. ten equal runners)	
	Probability (0–100%)	Surprise (1–5)	Probability (0–100%)	Surprise (1–5)
Surprise first	40.8	4.00	14.1	3.76
Surprise last	33.9	4.32	14.0	4.15
Mean	37.4	4.16	14.1	3.96

rated surprise before probability. The design thus was a 2 (single vs. multiple alternatives) × 2 (surprise ratings first vs. last).

2.2.2. Results

The probability assessments and surprise ratings are presented in Table 2. Most participants in the multiple-alternative condition correctly assumed that Erik's chances are low, in fact close to 1/11 (under the assumption that two laps before the finish, all 11 runners have an equal probability to win). The mean probability estimate in the single-alternative condition was substantially larger (37.4% against 14.1%). This large difference in probability was not matched by a corresponding difference in surprise: in fact, participants seem to be somewhat more surprised if Erik beats the stronger opponent than if he beats the group of equal opponents, though this difference was not statistically significant.

The results from Experiment 1 confirmed Attneave's conjecture that from a psychological perspective, surprise is not a simple function of the probability of an outcome, but "seems to depend on the probability of the event relative to the probabilities of other alternatives" (Attneave, 1959, p. 7). The track race scenario in Experiment 2 demonstrated that people can be equally or more surprised by a winner with $P > 0.30$ than with a winner with $P < 0.15$ (according to their own estimates), if the former has to outperform a single, stronger opponent.

3. Surprise associated with novel experiences

In the studies reported above, the surprisingness of the target outcome appeared to be due to its contrast with the probability of other potential outcomes. With one more probable alternative, the target seems to be more surprising than in a context of several, equally probable alternatives. We now present two situations in which there is a contrast between the target outcome and the protagonist's prior experience. Our hypothesis is that accumulated personal experience is more important in building up a person's dominant expectation than the judgmental processes leading to a probability estimate. Surprise has often been defined as a reaction to novelty (e.g. Berlyne, 1960), yet there appear to be no studies showing whether the surprise of novelty is primarily due to the low probability (or unexpectedness) of novel stimuli, or their contrast with the old and familiar. Following the contrast interpretation, an unfamiliar situation could be surprising even if it conforms to

verbally stated predictions. Experiment 3 shows that a person who never won before, but has a higher probability of winning, is believed to be more surprised than a more experienced (but less likely) winner. Experiment 4 demonstrates how a repeated outcome is believed to build up expectations of an outcome shift (in line with the gambler's fallacy), without reducing the surprise of such a shift. From the contrast hypothesis we predict that a changed outcome will be perceived to be more surprising after an unbroken series of identical outcomes, despite its increased probability.

3.1. Experiment 3

3.1.1. Method

3.1.1.1. Participants and design Norwegian university students attending a class of introductory psychology ($N = 52$) were randomly allocated to one of two conditions, by receiving different versions of the following scenario:

> Ellen and Nina have for the last few months bought instant lottery tickets about once a week. Ellen usually buys one or two scratch-tickets each time. A couple of months ago, she won NOK 500, and three weeks ago she won NOK 100. Nina usually buys three to five scratch-tickets each time, but has never won. This week Ellen buys two scratch-tickets and Nina buys five scratch-tickets.

Condition A: Before scratching their tickets they discuss their chances of winning. Who do you think will estimate her probability higher? Ellen/Nina/equal probability. *Condition B*: When scratching their tickets they discover that both have won NOK 1000 [approximately $110]. Who do you think will be more surprised? Ellen/Nina/equally surprised

3.1.2. Results

A majority of the participants in Condition A thought that Nina would estimate her probability higher than Ellen, in line with the normative view: the more tickets, the higher the chances. However, when asked who will be more surprised, participants in Condition B also favored Nina (Table 3, upper half). Both differences are significant ($P < 0.01$) by a binomial test. These surprise ratings are evidently not based on probability estimates, but must be due to the players' different histories of winning. A prize is regarded as more

Table 3
Number of choices (percentages) of more probable and more surprised winner in Experiments 3 and 4

Experiment 3: scratchcard scenario	Ellen	Nina	Equal	Total *n*
A: Higher probability	5 (19%)	16 (59%)	6 (22%)	27
B: Higher surprise	4 (16%)	13 (52%)	8 (32%)	25
Experiment 4: gambling scenario	**Peter**	**Henk**	**Equal**	**Total *n***
A: Higher probability	14 (21%)	32 (48%)	21 (31%)	67
B: Higher surprise	18 (29%)	26 (41%)	19 (30%)	63

surprising for a player who has never won than for a player with some previous experience of winning, despite the probabilities involved. This is compatible with the contrast interpretation of surprise: an outcome is seen as more surprising the more it contrasts with previous outcomes.

3.2. Experiment 4

3.2.1. Method

3.2.1.1. Participants and design A total of 130 students from the University of Utrecht and the Free University of Amsterdam were among other (unrelated) tasks presented with the following scenario:

> Peter and Henk are playing the roulette at a casino. Peter has lost three times in a row. He decides to give it a final try. Henk has lost five times in a row. He also decides to make a final try.

Half of the participants were randomly assigned to condition A, and the other half to Condition B, by receiving the following question: *Condition A*: Who, in your opinion, will assess his winning probability as higher: Peter or Henk? Peter/Henk/the same probability. *Condition B*: They both win. Who, in your opinion, will be more surprised: Peter or Henk? Peter/Henk/equally surprised

A separate group of 49 students, from the Free University of Amsterdam, were asked about their own (rather than the player's) probabilities and surprise. They were asked to answer one of the following two questions: (A) Who, in your opinion, has a higher probability of winning? Peter/Henk/equal ($n = 25$). (B) Given that both win, the winning of whom will surprise you more? Peter/Henk/equal ($n = 24$).

3.2.2. Results and discussion

As shown in Table 3 (lower half), a majority of participants in Condition A thought that Henk would assess his probability higher than Peter. The difference between Henk and Peter is significant by a binomial test ($P = 0.012$). This is in line with the classic gambler's fallacy (e.g. Keren & Lewis, 1994), according to which a player will think that a series of identical outcomes implies an increased probability for the complementary outcome. Yet, participants in Condition B did not think that Henk should be less surprised, presumably because the final outcome contrasts in his case to a series of six losses, whereas for Peter, the final win contrasts with three previous losses only. A comparison between those participants who distinguish between Peter and Henk (excluding those participants who considered their surprise or probability as equal) reveals a significant difference between the trends in the surprise and the probability conditions ($\chi^2(2, n = 90) = 7.48$, $P < 0.01$).

Finally, the pattern of results for the two additional groups who were asked to make their *own* judgments regarding the winning probabilities and associated surprise was different. Most participants (23, or 92%) thought that the winning probability of Peter and Henk was actually the same, and the majority (17, or 71%) indicated that they would be equally surprised by the winning of either protagonist. The fact that participants, in

these two control conditions, did not commit the gambler's fallacy, and said they would be equally surprised by both outcomes, suggests that they were able to take an objective, outside view. In contrast, participants in the experimental conditions, who were instructed to adopt an inside view, realized that Peter's and Henk's subjective situations differed from each other, implying expectations of different strength, as well as different amounts of surprise.

Both experiments in this section show that an outcome is perceived to be surprising when it contrasts with previous outcomes. In the scratch-ticket scenario, the player who never won before is considered to be more surprised than the player with some winning experience. In the gambling scenario, none of the players have won previously, but the player with a longer history of losing is rated as being more surprised than a player with a smaller string of losses. Participants rated in both cases the *players'* surprise.

Evidently, probability ratings and surprise ratings are in these scenarios based on different considerations. In the instant lottery scenario, the player with more tickets supposedly perceived her chances as being higher than those of the player with fewer tickets. This is in agreement with normative probability theory, but could also be due to the fact that she has never won before, and so in a sense more strongly "deserves" a positive outcome. It is, so to speak, "her turn". This line of reasoning was explored in the gambling scenario, where participants assumed that the player with the longer string of losses would think of his chances as being better than those of the player who had lost only three times in a row. Again, participants in a control condition, who were asked which player *they* thought had the higher chance, distanced themselves from the gambler's fallacy and maintained that Peter and Henk had the same probability of winning. The results from both scenarios demonstrate that high surprise and low subjective probability are not the same, as winning can appear at the same time, both quite probable and quite surprising.

4. Surprise of deviant outcomes

If surprise depends on a contrast between the obtained and the expected outcome, it could be affected by the difference between these two outcomes on dimensions other than the probabilities involved. An obtained outcome that differs widely from the expected outcome would increase the contrast, and hence appear more surprising than a more similar outcome or an outcome that is perceived to be closer to expectation. In many cases, similarity can also be used as an indication of probability. Indeed, the well-known *representativeness heuristic* (Kahneman & Tversky, 1972; Tversky & Kahneman, 1974) is based on the assumption that outcomes similar to the most typical or central instance of a category are perceived to be more likely than more dissimilar outcomes. For instance, in a series of coin tosses, four heads in a row appear less likely than a sequence consisting of two heads and two tails, the latter being perceived to be more "representative" for a random sequence of binomial outcomes. The contrast interpretation of surprise implies, however, that two events can differ in surprise even if they are perceived to be equally probable, if one of them appears to be more dissimilar or deviant from the most likely (expected) outcome on a prominent dimension. In the following studies this is obtained by presenting two competing alternatives that vary in degree of similarity to the expected

outcome. Following the contrast hypothesis, we predicted that the alternative that was most different from the expected outcome would be considered most surprising, even when the probability of the two competing outcomes was assumed to be the same.

4.1. Experiment 5

This experiment was designed to compare the surprise of two unexpected events with the same probability. The two differed only in their similarity, or closeness, to the expected event.

4.1.1. Method

4.1.1.1. Participants The participants were 76 students at the University of Utrecht (Group 1) who received the question of surprise together with several other, unrelated tasks, and 44 students at the University of Eindhoven (Group 2), who were individually approached on campus and asked to volunteer (unpaid) for a single question on surprise.

4.1.1.2. Design and procedure All participants were presented with the following scenario:

> Tom and Fred have followed a course in world history in college. The course is to be completed with an oral exam. The themes for this exam come from the course book in world history. This book contains one chapter on 'Prehistoric times', one on 'Ancient history', and one on the 'Middle ages'. The remaining 7 chapters cover 'Modern history' (after Columbus), so most questions will come from this part of the book. During the exam, the candidate draws one of ten cards corresponding to the book chapters, to determine the chapter on which he will be tested.

Tom and Fred expect both to draw a card with questions from one of the chapters on 'Modern history'. What, in your opinion, is the probability that they will draw a question from this part of the book?

The participants were then told, on a separate page, that actually Tom drew a card from the chapter on Prehistoric times, and Fred drew a card from the chapter on the Middle ages. Who would be more surprised? Participants in Group 1 had to choose one of the following three options: Tom/Fred/equally surprised. For participants in Group 2, the last option was removed so they had to strictly judge whether Tom or Fred was more surprised.

It was conjectured that a question from the Prehistory chapter would be perceived as further apart from the more likely type of question (Modern history), and thus create more surprise than a question from the Middle ages chapter, even though the probabilities of both questions were the same.

4.1.2. Results

Sixty-six of the 76 participants in Group 1 gave the 'correct' probability ($P = 0.7$) for a question from the Modern history part of the book. Of these, 51 said that Tom and Fred would be equally surprised, 12 said that Tom (Prehistory) would be more surprised, and

only three said that Fred, who drew the Middle ages chapter, would be more surprised. The difference between Tom and Fred is significant by a binomial test ($P < 0.05$, two-tailed).

In Group 2, 42 of 44 participants gave the correct probability. Of these, 29 thought the Prehistory question would be more surprising, against three who thought the Middle ages question would be more surprising. The difference is highly significant by a binomial test ($P < 0.001$). The remaining ten participants said that the questions were equally surprising (despite the fact that the questionnaire contained no "equal" option).

This study compared the perceived surprise of two equally unexpected exam questions. The similarity of the questions led most participants in Group 1 and even some in Group 2 to claim that the candidates would be equally surprised. More importantly, those who assumed that one candidate would be more surprised than the other believed in most of the cases that a question from the first chapter in the book would be more surprising than a question from the third chapter. The reason must be that the first chapter differs more (by its position as well as by its content) than the third chapter from the remaining, most expected section of the book.

4.2. Experiment 6

This experiment was designed to test the relationship between dissimilarity and surprise in yet another domain, where the problem of equal surprise was avoided by asking participants which of two objects would be selected as a more surprising gift. The design also controls for the possibility that one gift is inherently more surprising than another, by placing it in a surprising and a non-surprising context. In the example below, when Johan wants to *surprise* Karin by bringing her an alternative gift to his usual bottle of red wine, we predict that he will bring her a bunch of flowers rather than a bottle of white wine. This prediction is based on the supposition that a bunch of flowers differs more than white wine from the regular bottle of red wine, thus creating a larger contrast between the expected and actual gift. Similarly, in another condition, if he wants to surprise her by bringing an alternative to his usual bunch of red flowers, we predict that he will bring her a bottle of wine rather than flowers of a different color.

4.2.1. Method

4.2.1.1. Participants The participants were 96 students at the University of Utrecht recruited in the same way as participants in the previous experiments. They were randomly assigned to four different groups: two Experimental conditions, E1 and E2, and two Control conditions, C1 and C2.

4.2.1.2. Design and procedure Participants in Condition E1 were presented with the following scenario:

> Johan visits Karin every Saturday and usually brings her a bottle of red wine. On a few occasions he has brought her some white wine or a bunch of flowers, instead of the red wine. He knows that she enjoys these gifts, too.
> This Saturday, Johan says to himself: Today, I want to surprise Karin. Instead of the

usual bottle of red wine, I will bring her:
(a) A bottle of white wine
(b) A bunch of flowers

Participants in Condition E2 were told the same story, with the difference that Johan usually brings a bunch of red flowers and is now contemplating to surprise Karin with a bunch of white flowers or a bottle of wine. We predicted that flowers would be more surprising than wine in the first condition, whereas wine would be more surprising than flowers in the second, being more dissimilar from Johan's regular gift.

The control conditions were designed to test Karin's perceived expectations. In Control condition 1, the first part of the scenario was identical to the first paragraph of E1 described above, but with the following continuation:

This Saturday, Karin is talking to one of her friends about Johan. In the course of the conversation, she says that Johan is such a nice guy, who always brings something for Saturday. Normally, this is a bottle of red wine, or else, he will most probably bring:
(a) A bottle of white wine
(b) A bunch of flowers

In Control condition 2, Johan's gifts were exchanged to be comparable with those described in E2.

4.2.2. Results and discussion

When Johan normally brings Karin a bottle of red wine, 93.5% of the participants in Condition 1 think that in order to "surprise" her, he will choose a bunch of flowers rather than white wine (see Table 4). This choice cannot be attributed simply to the inherent 'surprisingness' of flowers. When Johan usually brings red flowers, a majority of 82.4% think that in order to surprise her, he will choose wine rather than flowers of a different color (Condition 2). Both these patterns of preference are significant with a binomial test ($P < 0.001$ and $P < 0.05$, respectively).

The results from the control conditions show that the most "surprising" gift (according to Johan) is not necessarily the least probable (according to Karin). In Control condition 1, only 38% of the participants think flowers are less probable than wine, the difference between the response patterns in the Experimental and Control conditions being highly significant ($\chi^2(1, N = 52) = 18.8$, $P < 0.001$). In Control condition 2, about half of the participants believe that Karin expects wine rather than flowers, demonstrating no clear

Table 4
Number of choices (percentages) of more surprising and more probable gift in four conditions of Experiment 6

	Wine	Flowers	Total *n*
E1: More surprising	2 (6%)	29 (94%)	31
C1: More probable	8 (38%)	13 (62%)	21
E2: More surprising	14 (82%)	3 (18%)	17
C2: More probable	14 (54%)	12 (46%)	26

pattern of expectancies, but resulting again in a significant difference between the Experimental and Control conditions ($\chi^2(1, N = 43) = 5.63$, $P < 0.02$). Evidently, the surprisingness of Johan's choices cannot be attributed to their improbability, or unexpectedness. If anything, Karin shows a tendency to expect the surprise. If Johan succeeds in surprising Karin, it must be due to the greater perceived contrast between the chosen and the regular gift.

5. General discussion

The studies presented in this article demonstrate that rated surprise and estimated (low) probabilities, although both reflecting aspects of a person's expectations, are not the same and can often differ. Low probabilities, in the present context, can be conceived as a necessary but not sufficient condition for surprise. It is the contextual background that will determine whether (and to what extent) low probabilities elicit surprise. Surprise appears to arise in situations in which a mental set, or an activated schema (Schuetzwohl, 1998) is interrupted by a novel, unexpected turn of events. The decisive factor is, accordingly, not just the low probability of the obtained outcome, but the relationship between the expected and the obtained.

Contrasts between expected and obtained outcomes that may lead to surprise can be achieved in different ways. The present studies explored surprises resulting from three different sources of contrast: contrasts due to relative probabilities, where an alternative outcome is more probable than the one actually occurring; contrasts formed by novelty and change, where the obtained outcome differs sharply from the individual's previous experience; and contrasts due to the perceptual or conceptual distance between the expected and the obtained. In all these cases, greater contrast is accompanied by higher ratings of surprise.

The contrast interpretation is certainly not limited to the three types explored in this paper, and is applicable to different situations. For instance, in a previous series of experiments, we found that positive and negative outcomes, with equal probability of occurrence, were not rated as equally surprising (Teigen & Keren, 2002). In these experiments, participants who were told about positive successful outcomes (e.g. absence of an allergic reaction; catching a flight; avoiding a parking ticket) rated the protagonist as clearly more surprised than those who were told about negative unsuccessful outcomes (allergic reaction; missed flight; parking ticket). This pattern of results appears to be inconsistent with prospect theory (Kahneman & Tversky, 1979), which posits that losses loom larger than gains, and hence could be expected to arouse a stronger emotional reaction.

A possible interpretation for the higher surprise associated with positive outcomes would be that people choose to prepare for the worst, to avoid disappointment and "unpleasant surprises". Indeed, disappointment can be minimized by lowering the probabilities for desired outcomes, by boosting probabilities of alternative undesired outcomes, or by imagining or mentally preparing for the unwanted. Such strategies are likely to make successes more surprising than equally probable failures.

Subsequent experiments, however, showed the opposite pattern of results in an achievement context. A student who expects to pass an exam, but fails, is considered to be more

surprised than a student who expects to fail, but passes. The difference between the surprising successes, in the first set of scenarios, and the surprising failures, in the achievement context, could be related to perceived outcome control. Positive outcomes due to chance appear to be more surprising than positive outcomes that are partly due to effort and ability.

These results make sense if we assume that people consider success as "normal" or "typical" in situations in which they have some degree of control, but not in out-of-control situations like lotteries and accidents. Winning appears to be less compatible with chance than losing. It is thus more surprising to win by chance than to lose by chance, even when the probabilities were stated to make the situations nominally equivalent. Evidently, formal equivalence is not enough; outcomes that contrast with people's intuitions about what they can "normally" expect evoke more surprise than warranted by the outcome probabilities involved.

The foregoing discussion clearly suggests a close link between surprise and disappointment. Indeed, disappointment, which has been defined as a "psychological reaction to an outcome that does not match up to expectations" (Bell, 1985, p. 1), is in many respects overlapping with the concept of surprise (applied to negative outcomes). As noted above, both can be reduced (or avoided) by imagining, and seriously considering, alternative pathways and corresponding outcomes. Whatever happens will then be less in contrast with one dominant frame of mind. The presence (or absence) of such thoughts or 'expectations' cannot be translated directly into probabilities. It is for instance completely sensible to prepare oneself mentally for several, mutually exclusive outcomes, but it is not normatively correct to make high-probability estimates of several alternative outcomes at the same time. Probabilities are, at least formally, constrained in the sense that a high probability of one outcome implies a low probability of its complement, and vice versa. There are no such rules for surprise, requiring a fixed amount of surprise to be distributed between all outcome alternatives.

In a similar vein, we propose that an obtained outcome will appear more surprising the more unfamiliar it is, and the less it is imagined and prepared for, compared to other outcomes. A person with long and varied experience will sometimes declare himself or herself immune against surprises, not meaning that everything is highly or even equally probable, but that he or she has witnessed and considered such a wide range of phenomena that no new incident will be too weird or too extreme not to be matched by those already encountered.

It is certainly not our claim, neither do our results suggest, that subjective probabilities and surprises are unrelated phenomena. Obviously they are both based on, and reflect, the presence or absence of expectations of varying strength. Changes in probabilities will, under equal circumstances, be accompanied by changes in surprise, and vice versa (Reisenzein, 2000). We have noticed that representativeness is one factor that seems to influence probability as well as surprise, and it may even be the case that estimated probabilities of unexpected outcomes are sometimes mediated by the amount of surprise they are predicted to arouse (Fisk & Pidgeon, 1998, 1997). Verbal probabilities may correspond better than numeric estimates to judgments of surprise, as in the case of the "Alternative-outcomes effect" (Windschitl & Wells, 1998), which shows that verbal phrases can be used to characterize the relative, rather than the absolute probability of

the target outcome. We predict that an even stronger alternative-outcomes effect could be obtained by asking people how surprised they would be, as indicated in Experiments 1 and 2 in the present article.

The studies reported in this paper have explored surprise by asking participants to rate other people's surprise in scenario situations, rather than their own surprise by real events. In a study of coincidences, Falk (1989) demonstrated that coincidences involving others were judged to be less surprising than similar coincidences happening to oneself. This study did not ask for probability estimates, so the possibility exists that people think that own coincidences are more improbable, perhaps because of one's own perceived uniqueness. But they may also be more surprised because events involving oneself are perceived as generally more impactful. Other studies indicate that the surprise value of an unexpected outcome can be influenced by its *importance*. Gendolla (1997) found that unexpected success or failure on an examination was rated as more surprising when the exam was important than when it was described as inconsequential for the student in question. In the present set of studies, the situations described were low in personal relevance for the participants, and hence potentially less surprising than real, personal events. Accordingly, the findings reported here reflect people's cognitive representations of surprise rather than their actual emotional experiences. This can be construed as a limitation in the "realism" of the present study, but at the same time it underscores, in our opinion, even more the discrepancy between surprise ratings and the probability estimates, as they are obtained under similar circumstances and could easily have been matched, unless our subjects had strong intuitions about their unique qualities. Surprise and probability ratings appear to offer distinctively different ways of assessing people's expectations. The questions "how probable is T", and "how surprising is T", may, as Attneave (1959) suggested, invite different answers.

Acknowledgements

This research was partly supported by grant no. 114193/330 from the Research Council of Norway to Karl H. Teigen, and by internal funds of Eindhoven University of Technology to Gideon Keren.

References

Attneave, F. (1959). *Applications of information theory to psychology: a summary of basic concepts, methods, and results*. New York: Holt, Rinehart and Winston.

Baillargeon, R., & Hanko-Summers, S. (1990). Is the top object adequately supported by the bottom object? *Cognitive Development*, *5*, 29–53.

Bell, D. E. (1985). Disappointment in decision making under uncertainty. *Operation Research*, *33*, 1–27.

Berlyne, D. E. (1960). *Conflict, arousal, and curiosity*. New York: McGraw-Hill.

Buck, R. (1999). The biological affects: a typology. *Psychological Review*, *106*, 301–336.

Choi, I., & Nisbett, R. E. (2000). Cultural psychology of surprise: holistic theories and recognition of contradictions. *Journal of Personality and Social Psychology*, *79*, 890–905.

Coughlan, R., & Connolly, R. (2001). Predicting affective responses to unexpected outcomes. *Organizational Behavior and Human Decision Processes*, *85*, 211–225.

Ekman, P., & Friesen, W. V. (1971). Constants across the cultures in the face and emotion. *Journal of Personality and Social Psychology*, *17*, 124–129.

Falk, R. (1989). The judgment of coincidences: mine versus yours. *American Journal of Psychology*, *102*, 477–493.

Fisk, J. E., & Pidgeon, N. (1997). The conjunction fallacy: the case for the existence of competing heuristic strategies. *British Journal of Psychology*, *88*, 1–27.

Fisk, J. E., & Pidgeon, N. (1998). Conditional probabilities, potential surprise, and the conjunction fallacy. *The Quarterly Journal of Experimental Psychology*, *51A*, 655–681.

Gendolla, G. H. E. (1997). Surprise in the context of achievement: the role of outcome valence and importance. *Motivation and Emotion*, *21*, 165–193.

Kahneman, D., & Tversky, A. (1972). Subjective probability: a judgment of representativeness. *Cognitive Psychology*, *3*, 430–454.

Kahneman, D., & Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica*, *47*, 263–291.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.

Keren, G., & Lewis, C. (1994). The two fallacies of gamblers: type I and type II. *Organizational Behavior and Human Decision Processes*, *60*, 75–89.

Mareschal, D., Plunkett, K., & Harris, P. (1999). A computational and neuropsychological account of object-oriented behaviours in infancy. *Developmental Science*, *2*, 306–317.

Mellers, B., Schwartz, A., & Ritov, I. (1999). Emotion-based choice. *Journal of Experimental Psychology: General*, *128*, 332–345.

Meyer, W. U., Niepel, M., Rudolph, U., & Schuetzwohl, A. (1992). An experimental analysis of surprise. *Cognition and Emotion*, *5*, 295–311.

Miller, D. T., Turnbull, W., & McFarland, C. (1990). Counterfactual thinking and social perception: thinking about what might have been. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (pp. 305–331), Vol. 23. New York: Academic Press.

Petty, R. E., Fleming, M. A., Priester, J. R., & Feinstein, A. H. (2001). Individual versus group interest violation: surprise as a determinant of argument scrutiny and persuasion. *Social Cognition*, *19*, 418–442.

Plutchik, R. (1991). *The emotions*. (Rev. ed.). New York: University Press of America.

Reisenzein, R. (2000). Exploring the strength of association between the components of emotion syndromes: the case of surprise. *Cognition and Emotion*, *14*, 1–38.

Roseman, I. J., Antoniou, A. A., & Jose, P. E. (1996). Appraisal determinants of emotions: constructing a more accurate and comprehensive theory. *Cognition and Emotion*, *10*, 241–277.

Schuetzwohl, A. (1998). Surprise and schema strength. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1182–1199.

Shackle, G. L. S. (1969). *Decisions, order and time in human affairs*. Cambridge: Cambridge University Press.

Slovic, P., & Fischhoff, B. (1977). On the psychology of experimental surprises. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 544–551.

Smedslund, J. (2002). From hypothesis-testing psychology to procedure-testing psychologic. *Review of General Psychology*, *6*, 51–72.

Teigen, K. H. (1988). When are low-probability events judged to be 'probable'? Effects of outcome-set characteristics on verbal probability estimates. *Acta Psychologica*, *67*, 157–174.

Teigen, K. H. (2001). Equal chances = good chances. Verbal probabilities and the equiprobability effect. *Organizational Behavior and Human Decision Processes*, *85*, 77–108.

Teigen, K. H., & Keren, G. (2002). When are successes more surprising than failures. *Cognition and Emotion*, *16*, 245–268.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science*, *185*, 1124–1131.

Windschitl, P. D., & Wells, G. L. (1998). The alternative-outcomes effect. *Journal of Personality and Social Psychology*, *75*, 1411–1423.

Windschitl, P. D., & Young, M. E. (2001). The influence of alternative outcomes on gut-level perceptions of certainty. *Organizational Behavior and Human Decision Processes*, *85*, 109–134.

Young, A. W., Rowland, D., Calder, A. J., Etcoff, N. L., Seth, A., & Perrett, D. I. (1997). Facial expression megamix: tests of dimensional and category accounts of emotion recognition. *Cognition*, *63*, 271–313.