# Getting Scarred and Winning Lotteries: Effects of Exemplar Cuing and Statistical Format on Imagining Low-Probability Events[†]

BEN R. NEWELL*, CHRIS J. MITCHELL and BRETT K. HAYES
*School of Psychology, University of New South Wales, Sydney, Australia*

ABSTRACT

Three experiments tested the exemplar cuing and frequency format accounts of how the 'imaginability' of low-probability events is enhanced. The experiments manipulated imaginability by varying the statistics used to describe negative (e.g. being scarred as a result of laser surgery) and positive (e.g. winning a lottery) low-probability events. The results strongly supported the frequency format account, whereby imaginability is enhanced through the use of frequency formats for conveying statistical information (e.g. 20 out of 2000 as opposed to 0.01%). However, only limited support was found for exemplar cuing (EC) theory. Overall the results support the claim that the imaginability of outcomes plays a key role in thinking about low-probability events, but question the mechanisms specified by EC theory for mediating such effects. Copyright © 2007 John Wiley & Sons, Ltd.

KEY WORDS    imaginability; exemplar cuing; frequency format; probability judgment

## INTRODUCTION

A fatal shark attack on a 17-year-old boy of West Beach in Adelaide left the popular beaches of South Australia almost deserted in the summer of 2004. Despite the extreme low probability of a shark attack occurring, it seems that the extensive media coverage following the incident was enough to evoke vivid, terrifying images of the attack in the minds of beach-goers, ensuring that they stayed away. Such an effect can be explained in terms of the 'availability heuristic'—the idea that people judge the probability of an event by the ease with which instances or associations can be brought to mind (Tversky & Kahneman, 1973).

One factor claimed to influence the availability of instances is whether statistical information is presented in terms of frequencies or in terms of probabilities (e.g. Finucane, Alkhami, Slovic, & Johnson, 2000;

---

*Correspondence to: Ben R. Newell, School of Psychology, University of New South Wales, Sydney 2052, Australia.
E-mail: ben.newell@unsw.edu.au
[†]Experiments 1 and 2 were presented at the 2005 Meeting of the Cognitive Science Society, Stresa, Italy.

Finucane, Peters, & Slovic, 2003; Loewenstein, Weber, Hsee, & Welch, 2001; Slovic, Finucane, Peters, & MacGregor, 2002). For example, Slovic, Monahan, and MacGregor (2000) demonstrated that clinicians provided with recidivism risks presented as frequencies (e.g. 20 out of 100) judged mental patients as posing higher risks than when the same information was presented as a probability (e.g. 20%).[1] The explanation was that only the frequency presentation generated a 'terrifying image' of the recidivist in the mind of the clinician and that the affect associated with this imagery led to the more extreme judgments (Slovic et al., 2000). In a similar vein, Purchase and Slovic (1999) reported that individuals were more frightened by information about chemical spills framed as frequencies (e.g. out of 1 000 000 exposed people, there will be 1 additional cancer death) than as probabilities (e.g. each exposed individual has an additional chance of 0.0001% of getting cancer). In a related set of findings Yamagishi (1997) demonstrated that participants rated a disease that kills 1286 people out of every 10 000 as more dangerous than one that kills 24.14% of the population.

In an intriguing article, Koehler and Macchi (2004) speculated that particular statistical formats need not necessarily evoke terrifying or affectively rich imagery to influence probability judgment; it may be sufficient for the statistics to simply evoke thoughts about other examples of the target event. Their 'exemplar cuing (EC) theory' states that, 'the weight decision makers attach to low-probability events is, in part, a function of whether they can easily generate or imagine exemplars for the event' (p. 540). They suggest, for example, that a lottery ticket might be more appealing if a potential purchaser is induced to think about other winning lottery tickets.

According to EC, however, the use of a frequency format is not the crucial factor underlying the imaginability of exemplars. Koehler and Macchi discuss two separate mechanisms for the facilitation of exemplar generation. The first—'multiplicative'—mechanism cues exemplars when the product of the size of the reference class for the event and the incidence rate of the event is greater than 1. For example, a lottery ticket described as having 1% chance of winning, with a reference class of 500 000 tickets sold in a day, generates 5000 exemplars of winning tickets. Such a ticket is deemed to be more appealing than a ticket with a 1% chance in a lottery in which only 50 tickets are sold in a day because this arrangement only generates 0.5 of a winning ticket. Importantly, this mechanism is unaffected by the format of the information, that is by whether incidence rates are provided as percentages (1%) or frequencies (1 out of 100). The reason that format does not affect exemplar generation is that the format does not identify a relevant sample space within which to search for exemplars (Koehler & Macchi, 2004). This sample space is provided by the reference class (e.g. the number of other tickets sold), a factor common to both formats. Macchi (2000) reported results supporting this claim, demonstrating that when the size of the reference class was controlled, the format of statistical information had no effect on probability judgment.

The second mechanism described by EC theory is the 'numerator' mechanism. This operates only on information presented as frequencies and thus its predictions overlap with those of the format account. Specifically the mechanism states that exemplars are cued when the numerator of the incidence rate is one or more (e.g. 20 out of 2000) but not when it is less than one (e.g. 0.2 out of 20). The reasoning is simple: only when the numerator is greater than one can examples of other 'winners' (in the lottery example) be generated (Koehler, 2001; Koehler & Macchi, 2004). This mechanism is derived from research on the ratio-bias that has indicated peoples' preferences for playing lotteries with large ratios over small ratios even when the two are statistically identical (e.g. preferring a lottery with 10 winning jelly beans out of 100 over one with 1 winning bean out of 10) (Denes-Raj & Epstein, 1994; Epstein, 1994). Koehler and Macchi (2004) follow previous research in explaining this tendency as arising because people find larger ratios more difficult to comprehend

---

[1]Although 20% is not strictly a probability (i.e. a number between 0 and 1), in keeping with previous literature on the probability/frequency distinction we use percentages to display the 'probability' of an event occurring.

than smaller ones, and people focus on the absolute number of ways an outcome can occur rather than on the probability ratio (Denes-Raj & Epstein, 1994; Miller, McFarland, & Turnbull, 1989).

Koehler and Macchi (2004) tested these predictions in the context of DNA evidence in mock jury experiments. A key issue with DNA evidence is the degree to which DNA matches can be attributed to mere coincidence (Koehler, 1997). Koehler and Macchi reasoned that if exemplars could be cued of innocent individuals who nonetheless had matching DNA, then the DNA evidence against the defendant would be given less weight.

In Koehler and Macchi's first experiment participants were presented with the incidences of co-incidental matches in either probability (e.g. 0.001%) or frequency formats (e.g. 1 in 100 000). Information about the 'problem-relevant sample space' or reference class was manipulated by telling participants that police believed the murderer lived in a town with either a large (5 000 000) or small population (500). The product of the incidence rate and the reference class was only greater than one for the 'large population' group (large: $5\,000\,000 \times 0.001\% = 50$; small: $500 \times 0.001\% = 0.005$), thus, Koehler and Macchi argued that exemplars of other individuals with DNA that matched the defendant's merely by co-incidence would only be cued in the large population group. The data were consistent with this prediction: participants thought the evidence against the defendant was weaker when the 50 exemplars of other possible matches were cued. Importantly, the format manipulation, whether the numbers were presented in frequencies or probabilities, did not have a significant effect on judgments.

In a second experiment, no mention was made of the reference class, but the numerator of the incidence rate statistic for the DNA match was presented as either a fraction or an integer. EC theory predicted that the integer (e.g. 1 out of 1000) and not the fraction (e.g. 0.1 out of 100) would call a juror's attention to other individuals who may match, despite the ratios being mathematically identical. Again, this prediction was supported: jurors were relatively less impressed by the evidence when a fractional numerator was used.

Koehler and Macchi's results are compelling and they raise a number of interesting issues. First, the claim that format has no effect on judgments when the 'multiplicative mechanism' is invoked is controversial given the wealth of evidence suggesting that frequency formats are crucial to enhancing imaginability (e.g. Finucane et al., 2003; Loewenstein et al., 2001; Purchase & Slovic, 1999; Slovic et al., 2000; Yamagishi, 1997). Second, evidence for the proposed multiplicative mechanism comes from one experiment, using one particular set of incidence rates and reference classes, and in one context (DNA evidence). The predictions of EC theory and the strong, 'algebraic' predictions of the multiplicative mechanism in particular, warrant further testing in different contexts in order to assess the generality of EC as a theory of how people think about low-probability events.

Koehler and Macchi (2004) speculated that the effects they observed in the DNA context could have wide ranging practical implications for a variety of situations in which information about low-probability events is conveyed. For example, would the number of exemplars generated affect people's willingness to take particular vaccines, or undergo medical treatments? Would participation in lotteries be affected by the extent to which exemplars of 'winners' could be cued? To the extent that EC is a general theory of how people think about low-probability events, one would expect to see exemplar-cuing effects in such scenarios that are consonant with those observed by Koehler and Macchi (2004).

A third, more general, issue, is the extent to which different formats do indeed affect imaginability. In most studies these effects are simply inferred from the responses on other dependent measures (e.g. willingness to take risks) (Slovic et al., 2002). Some work has been done examining the relation between imagery and decision-making (e.g. Slovic, Layman, Kraus, Flynn, Chalmers, & Gesell, 1991) but evidence of the direct link between frequency formats and enhanced imaginability is currently lacking in the literature. In our third experiment, we sought evidence that format manipulations influence imaginability directly.

Thus, the goals of our three experiments were to test (1) the specific mechanisms proposed by the EC account, (2) the contrasting predictions of the frequency and EC accounts regarding the format effect, (3) the

extent to which the effects hold across different contexts, and (4) for direct effects of format on ratings of imaginability.

## OVERVIEW OF EXPERIMENTS

The three experiments used a design which aimed to test the contrasting predictions of EC theory and the frequency format account. The design was as follows: Question format (probability or frequency) was manipulated within subjects (in Experiments 1 and 2, between subjects in Experiment 3), and number of exemplars cued (less than one or greater than one) was manipulated between subjects in all three experiments. The multiplicative mechanism of the EC account was tested by comparing responses to probability format questions in which the product of the reference class and the incidence rate was more than one (e.g. 2000 and 1%), with questions in which the product was fewer than one (e.g. 20 and 1%). The numerator mechanism was tested by comparing responses to frequency format questions in which the numerator was greater than one (e.g. 50) to those in which it was less than one (e.g. 0.5).

In this design, the EC account predicts a main effect of number of exemplars cued such that people should judge an event as more likely to occur when more than one exemplar is cued either via the multiplicative or numerator mechanisms. However, the account predicts no effect of question format. In contrast, the frequency account predicts a main effect of question format such that events should be judged more likely when a frequency format is used. It also predicts an effect of the number of exemplars cued, but, crucially, only in the frequency format condition. This is because in the probability format no vivid images are promoted (regardless of the product of the incidence rate and reference class) whereas in the frequency format, the integer numerator (e.g. 20) is more likely to result in 'vivid imagery' than the fractional one (e.g. 0.2) (cf. Epstein, 1994; Koehler, 2001; Slovic et al., 2000; Yamagishi, 1997). This latter prediction is, of course, identical to that of the numerator mechanism of EC theory.

### Experiment 1
Experiment 1 used two scenarios: willingness to participate in a lottery and willingness to call a TV station for the chance of participating in a game show. In the case of lottery, the low-probability event was buying a winning ticket; for the game show it was appearing on the show. EC theory predicts that participants should be more willing to buy a ticket in the lottery or call the TV station if either the numerator or the multiplicative mechanism operates to generate exemplars of other winning lottery tickets or successful game show entrants. The frequency account predicts that willingness should increase when frequency formats (which are claimed to facilitate imaginability) are used.

### Method

*Participants.* Eighty undergraduate students from the University of New South Wales participated in the experiment.

*Design and procedure.* The experiment was a 2 (number of exemplars cued: $>1$, $<1$) $\times$ (2) (format: probability, frequency) mixed design with repeated measures on the second factor. Participants were provided with a questionnaire containing two questions about willingness to buy a lottery ticket and to participate in a TV game show. One question was in a frequency format about one of the scenarios, the other in a probability format about the other scenario. For one group both these questions cued exemplars, for the

Table 1. Design of Experiments 1, 2 and 3

| No. of exemplars cued | Question format | |
|---|---|---|
| | Probability | Frequency |
| >1 | 1%, 5000 (50) | 50, 5000 (50) |
| <1 | 1%, 50 (0) | 0.5, 50 (0) |

*Note*: Number of exemplars cued (according to EC Theory) for each question type is given in parentheses. Number of exemplars cued was manipulated between participants and question format was manipulated within subjects in Experiments 1 and 2 and between subjects in Experiment 3.

other group neither of the questions cued exemplars. The order of frequency/probability questions and the assignment of formats to the scenarios were fully counterbalanced.

In the lottery scenario, the reference class referred to the number of tickets sold in a day (50 or 5000); in the game show scenario, the reference class was the number of people calling the TV station each day (20 or 2000). The incidence rate of the positive event (lottery: buying a winning ticket; game show: getting on the show) was given as a probability (1%) or as the equivalent raw frequency (e.g. 50 out of 5000 or 0.5 out of 50).

Table 1 displays the design of Experiment 1 with the numbers used in the lottery scenario questions given as examples. To give a concrete example: a participant in the 'less than one exemplar cued' condition was asked one probability question about buying a ticket with a 1% chance of winning in a lottery in which 50 tickets are sold each day; and one frequency question about calling a TV show in which 0.2 of every 20 callers are selected and 20 people call the station each day (The Appendix provides examples of the actual questions used.). After reading each scenario, participants made a 'willingness to buy a ticket/call the TV station' response on a 7-point Likert scale anchored by 1—'Not at all willing' and 7—'Very willing'.

*Results and discussion*

Preliminary analysis indicated that there were no significant effects of scenario type (lottery vs. game show, largest $F = 1.8$) and so the data were collapsed across the two in all subsequent analyses. Figure 1 displays the mean willingness ratings collapsed across the lottery and game show scenarios. Visual inspection of the figure suggests increased willingness when the number of exemplars cued were fewer than one (the bars on the right are higher than those on the left), and when a frequency format was used (the open bars are higher than the filled bars).

Statistical analyses confirmed this pattern. A 2 (number of exemplars cued: >1, <1) × (2) (format: probability, frequency) mixed ANOVA with repeated measures on the second factor showed a main effect of number of exemplars cued $F(1, 78) = 4.21$, $p < .05$, in the opposite direction to that predicted by EC theory. Participants were significantly more willing to engage in the behaviour (buy a lottery ticket or participate in a game show) when (according to the theory) fewer than 1 exemplar of the low-probability event occurring was cued. There was also a main effect of format, $F(1, 78) = 10.05$, $p = .002$, indicating greater willingness to engage in the behaviour when frequency formats were used. The interaction was not significant ($F < 1$). One other interesting aspect of the data in the frequency format conditions is that there appeared to be a tendency for participants to be more willing to participate in the behaviours when a fractional numerator was used than when an integer was used. Although, this difference was not statistically significant ($F(1, 79) = 1.78, p > .01$), it is in the opposite direction to the prediction of EC's numerator mechanism.

The results of Experiment 1 provide clear support for the format account—willingness to participate increased with the use of frequency formats. However, the results offer little support for EC theory. According to the theory, no exemplars of other winning tickets or successful game show entrants should have been cued
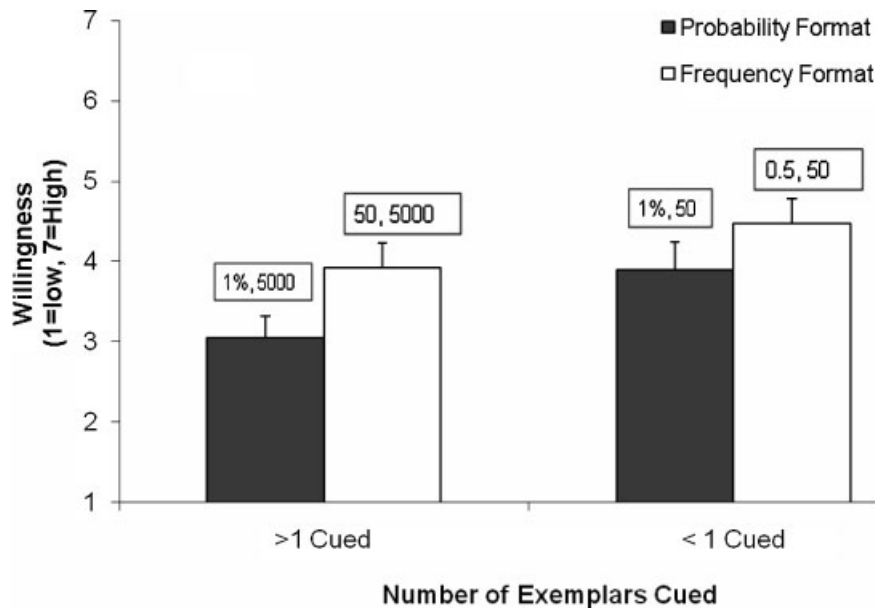
Figure 1. Mean willingness ratings collapsed across the two positive scenarios in Experiment 1. The reference classes and incidence rates used in the lottery scenario are displayed as an illustration of the statistical format used for each condition. Error bars are standard error of the mean

in the two <1 conditions, but participants were significantly more willing to engage in the behaviours in those conditions than in the >1 conditions. Furthermore, the main effect showing higher willingness overall in frequency format conditions cannot be accounted for by EC theory (Koehler & Macchi, 2004; Macchi, 2000) because the reference class is equated across the format conditions. We will return to a more detailed examination of these issues in the General Discussion Section but first we turn to one important difference between the scenarios used in this experiment and the one used by Koehler and Macchi (2004). In their scenario the imagined event was the wrongful conviction of a murder suspect—a negative outcome—in comparison our imagined events are positive—winning lotteries or appearing on a game show. Although EC theory does not appear to make different predictions regarding positive and negative events, it is possible that the role played by a cued exemplar differs depending on whether the imagined event is desirable. We examined this possibility in Experiment 2 by using scenarios with negative events.

## Experiment 2

Experiment 2 tested the predictions of EC theory and the format account in scenarios in which the low-probability event was a negative outcome for the participant. Participants were asked to indicate their willingness to take part in clinical trials of a new laser treatment for birthmarks and a new flu vaccine. In the case of the laser treatment, the low-probability event was suffering from permanent scarring as a result of having the treatment. For the vaccine, it was suffering from an unpleasant allergic reaction following administration of the vaccine. EC theory predicts that if either cuing mechanism operates to generate exemplars of other people who have suffered these side effects, participants should be less willing to participate in clinical trials of the treatments. In contrast the format account predicts that willingness should decrease when a frequency format is used but only predicts an effect of number of exemplars cued in the frequency format condition.

As in Experiment 1, we manipulated the number of exemplars cued (according to EC) by (a) varying the number of individuals involved in the trials (i.e. reference class), and (b) varying the numerator of the incidence rate.

### Method

*Participants.* One hundred and forty undergraduate students from the University of New South Wales participated in the experiment.

*Design and procedure.* The design was identical to that of Experiment 1 with the exception that participants were given negative event scenarios: willingness to participate in clinical trials of a flu vaccine and a laser treatment for birthmark removal. The reference class sizes, incidence rates and numerators were manipulated in the same way as in Experiment 1. All numerical values were identical. The size of the reference class was manipulated by indicating how many other people would participate in the trial (vaccine: 20 or 2000; laser: 50 or 5000). Participants used a 7-point Likert scale to rate their willingness to participate in the clinical trials.

### Results and discussion

Preliminary analysis indicated that there were no significant effects of scenario type (lottery vs. game show, largest $F = 1.9$) and so the data were collapsed across the two in all subsequent analyses.

Figure 2 displays the mean willingness ratings collapsed across the flu vaccine and laser treatment scenarios. Inspection of the figure suggests no effect of the number of exemplars cued on willingness (the average of the bars on the left and right is similar), but lower willingness in the frequency format condition (the open bars are lower than the filled bars).

Statistical analyses confirmed this impression. A 2 (number of exemplars cued: $>1$, $<1$) $\times$ (2) (format: probability, frequency) mixed ANOVA with repeated measures on the second factor showed no main effect of number of exemplars cued ($F < 1$), but a significant main effect of format, $F(1, 138) = 12.57$, $p < .001$. The
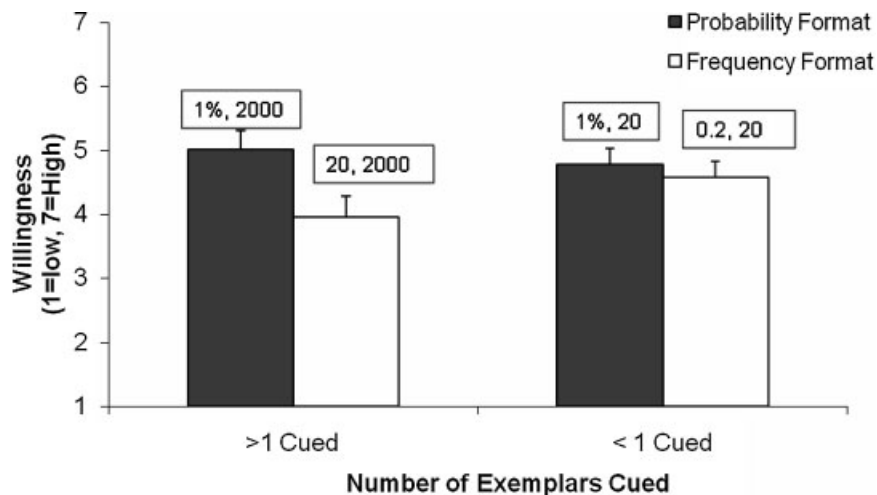


Figure 2. Mean willingness ratings collapsed across the two negative scenarios in Experiment 2. The reference classes and incidence rates used in the vaccine scenario are displayed as an illustration of the statistical format used for each condition. Error bars are standard error of the mean

format effect indicates a lesser willingness to participate in the clinical trials when a frequency format was used. However, this effect was modulated by an interaction between format and number of exemplars cued, $F(1, 138) = 5.84, p < .05$. Simple effects analysis revealed that the number of exemplars cued had an effect on willingness with the frequency format, $F(1, 139) = 4.19, p < .05$, but not with the probability format ($F < 1$). That is, a frequency format led participants to be significantly less willing to participate in the clinical trials when an integer numerator (e.g. 50) was used than when a fractional one (e.g. 0.5) was used.

EC theory predicted a main effect of number of exemplars cued but this effect was only present in the frequency format condition. This pattern of results provides support for the 'numerator' mechanism of EC theory but not for the 'multiplicative' mechanism. Notably though, the numerator effect is also predicted by the format account, and this account also predicts the observed main effect of question format. Given that the latter effect is not predicted by EC theory, and that no main effect of number of exemplars cued was observed, the overall pattern of results appears to provide much better support for the frequency format account (Finucane et al., 2003; Loewenstein et al., 2001; Slovic et al., 2000).

As noted in the Introduction Section, format effects of the kind found in Experiments 1 and 2 are often attributed to the facilitation of imaginability under frequency formats. However, the evidence for this underlying mechanism is indirect. Our results (and those reported previously) typically provide no independent evidence that participants given frequency formats did indeed experience more vivid imagery than those given probability formats. In Experiment 3, we sought more direct evidence that ease of imagining would be more strongly associated with frequency than probability formats.

## Experiment 3

In Experiment 3, participants were presented with both the positive scenarios tested in Experiment 1 (lottery and game show) and the negative ones from Experiment 2 (vaccine and laser treatment). This allowed us to examine differences in responding to positive and negative scenarios in the same participants and in a single experiment. The design was similar to that used in Experiments 1 and 2, but in addition to eliciting 'willingness ratings', we asked participants for ratings of the imaginability of scenarios (e.g. 'How easy is it to imagine yourself being a winner in the lottery?'). This was done in an attempt to find clearer evidence of the potential mechanisms underlying the format effects found in Experiments 1 and 2. Standard interpretations of the frequency format account would predict that ratings of imaginability will be higher for frequency than probability formats. We also asked participants for emotion ratings (e.g. 'How do you feel when you imagine being a winner of this lottery?'); these served as a 'manipulation check' to ensure that our negative and positive scenarios elicited different emotional reactions.

### Method

*Participants.*  One hundred undergraduate students from the University of New South Wales participated in the experiment in return for course credit.

*Design.*  The experiment was a 2 (number of exemplars cued: >1, <1) × 2 (format: probability, frequency) between subjects design. In contrast to Experiments 1 and 2, format was manipulated between rather than within subjects. This change was made due to the concern that participants might become aware of the intention of the format manipulation if asked for willingness and imaginability ratings in both formats for all four scenarios (cf. Kahneman, 2003).

*Materials.*  The scenarios used in Experiments 1 and 2 were used in this experiment. The numbers used for the reference classes and incidence rates remained the same. However, one slight change was made to the

positive scenarios in which the chance of winning the lottery (or gaining entry to the game show) was less than one. It is possible that some participants may have been confused by these statistics, reasoning that there should always be at least *one* winner or *one* person selected. To counter this possibility, these scenarios featured an additional instruction—'Please note that it is not guaranteed that there will always be a winner'—which appeared below the scenario.

*Procedure.* Participants were assigned randomly to one of the four conditions obtained by crossing the format and number of exemplars cued factors. Participants completed the experiment individually on a computer. Each participant was presented with four scenarios, two positive (lottery and game show), and two negative (vaccine and laser). For each participant all four scenarios were in either probability or frequency format and either promoted the cuing of more than one exemplar or did not. The order of scenarios was randomised for each participant.

A single scenario appeared on the screen; participants read it and then answered three questions: (1) willingness to participate, (2) ease of imaginability and (3) emotion/feeling rating. The wording of the imaginability and emotion/feeling questions was as follows: lottery scenario (imaginability)—'How easy is it to imagine yourself being a winner in the lottery?', (emotion) 'How do you feel when you imagine being a winner of this lottery?'; game show (imaginability)—'How easy is it to imagine yourself being selected for the game show?', (emotion) 'How do you feel when you imagine being selected for this game show?'; vaccine (imaginability)—'How easy is it to imagine yourself suffering an unpleasant allergic reaction to the vaccine?', (emotion) 'How do you feel when you imagine suffering an unpleasant allergic reaction to this vaccine?'; laser (imaginability)—'How easy is it to imagine yourself suffering permanent scarring from the laser treatment?', (emotion) 'How do you feel when you imagine yourself suffering permanent scarring by the treatment?'

The willingness, imaginability and emotion questions were always presented in the same order and answers were made by adjusting an on-screen slider control. The answer to each question disappeared from the screen after a rating was made. Willingness ratings were elicited using the same scale as in Experiments 1 and 2 (1 = not at all willing, 7 = very willing); imaginability ratings were elicited on a scale anchored by (1 = very difficult to imagine and 7 = very easy to imagine); emotion ratings were elicited on a scale anchored by (1 = very negative emotion (sad, scared), 4 = neutral and 7 = very positive emotion (happy, excited)).

After answering the three questions for all four scenarios, participants completed a 10-item numeracy questionnaire (Lipkus, Samser, & Rimer, 2001). This was included to examine whether differences in numerical ability might explain any observed differences in willingness and imaginability ratings. All participants performed well on this measure (mean = 85.6%, $SD = 14.1$) and it did not correlate with or predict any aspect of performance and so is not discussed further.

### Results and discussion

*Manipulation check: Emotion ratings.* In Experiments 1 and 2, we suggested that the scenarios we used were differentiated in terms of the type of emotion evoked—positive versus negative—but we had no independent evidence of this difference. The emotion rating scores served as a manipulation check to examine whether our scenarios were indeed perceived differently. Higher numbers on the 7-point scale indicated more positive emotions (excited, happy) and lower numbers more negative emotions (sad, scared). The mean ratings for the positive scenarios (collapsed across format and number of exemplars cued) were, lottery = 6.23 ($SD = 1.11$), game show = 5.01 ($SD = 1.56$) and for the negative scenarios were, vaccine = 2.35 ($SD = 0.96$) and laser = 1.59 ($SD = 0.73$). Collapsing across scenarios and comparing positive versus negative revealed a clear difference ($F(1, 99) = 639.4$, $p < .001$) indicating that the scenarios were differentiated in terms of perceived emotional content. Comparisons of scenarios *within* the two emotion types also revealed significant

differences in the strength of perceived emotions. In the positive scenarios, the lottery scenario was rated as more positive than the game show, $F(1, 99) = 73.5$, $p < .001$; and in the negative scenarios the laser scenario was rated as more negative than the vaccine, $F(1, 99) = 60.9$, $p < .001$.

These data support the claim that the reversal in the patterns of willingness found in Experiment 1 (more willing to participate when frequency formats were used) and Experiment 2 (less willing to participate when frequency formats were used) was driven by the difference in the emotional valence of the scenarios—positive and negative, respectively. It follows from this that the willingness effects found in Experiments 1 and 2 are likely to be replicated more reliably in those scenarios that elicited the strongest emotional reactions (i.e. lottery and laser). To examine this possibility, and for ease of comparison across scenarios, the results for the willingness and imaginability ratings are presented for individual scenarios rather than collapsed across the two positive and two negative scenarios.

*Willingness ratings.* Figure 3 displays the mean willingness ratings for the lottery, game show, vaccine and laser scenarios, respectively. Taking the two positive scenarios first, the figure suggests an increased willingness when a frequency format was used (the average of the lighter bars is higher than the dark bars), though the effect is clearer in the scenario which evoked the stronger positive emotion (i.e. the lottery scenario). There appears to be no systematic effect of number of exemplars cued (the two bars on the left of each group of four are not consistently higher or lower than the two on the right). Statistical analyses confirmed this impression. The effect of format was significant only in the lottery scenario, $F(1, 96) = 4.01$, $p < .05$ (game show $F < 1$) and there was no effect of exemplars cued in either scenario ($Fs < 1$). The interaction was not significant in either scenario ($Fs < 1$).

The increased willingness to participate when frequency formats are used is consistent with the effect found in Experiment 1. The novel finding is that this effect appears more reliable when the emotional reaction is stronger. The 'contradictory' number of exemplars cued effect found in Experiment 1—that is, greater willingness when *fewer* than one exemplar was cued—was not replicated. It is possible that willingness was reduced in the <1 conditions by the addition of the instruction that it was not 'guaranteed that there will always be a winner'. Despite the absence of this reversed exemplar cued effect, there was no indication of an effect consistent with the EC theory prediction (i.e. greater willingness in the >1 conditions).
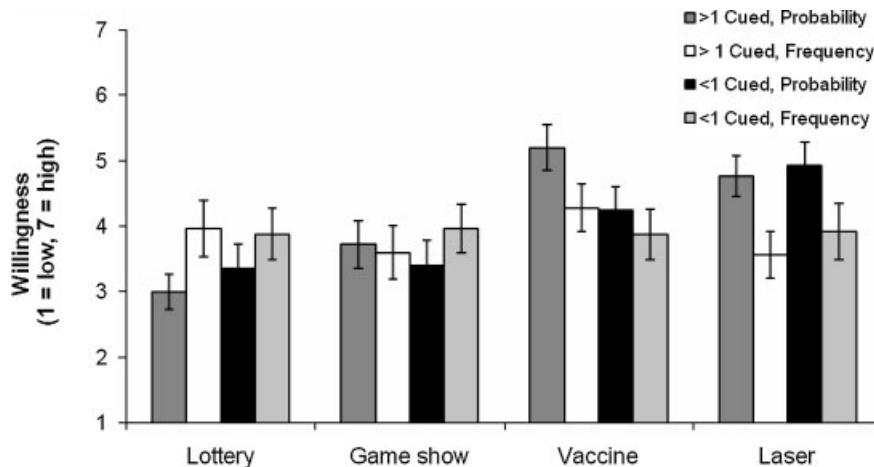


Figure 3. Mean willingness ratings for each scenario in Experiment 3. Although both factors were manipulated between subjects, bars are presented grouped for each scenario to facilitate cross-scenario comparisons. Error bars are standard error of the mean

Ratings from the negative scenarios indicate decreased willingness when frequency formats were used (lighter bars are lower than the darker bars) in both scenarios. Consistent with the positive scenarios, the format effect appears more reliable in the scenario which evoked the stronger emotional reaction (i.e. the laser scenario). There were no systematic effects of exemplars cued in the laser scenario, but a tendency toward greater willingness when more than one exemplar was cued in the vaccine scenario (the two bars on the left are higher than those on the right in the vaccine scenario). Statistical analyses confirmed an effect of format in the laser scenario $F(1, 96) = 9.00$, $p < .01$; and a marginally significant effect of format in the vaccine scenario $F(1, 96) = 3.07$, $p = .08$. Number of exemplars cued had no effect in the laser scenario ($F < 1$) and a marginally significant effect in the vaccine scenario $F(1, 96) = 3.46$, $p = .07$. The interaction was not significant in either scenario ($Fs < 1$).

The decreased willingness to participate when frequency formats are used is consistent with the effect found in Experiment 2. Again, the novel finding is that this effect is more reliable when emotional reactions are perceived more strongly. The effect of number of exemplars cued is not as clear-cut. In the laser scenario there was some indication of the numerator mechanism in operation (i.e. greater willingness in the <1 frequency condition than the >1 frequency condition) but the effect was smaller than that found in Experiment 2. In the vaccine scenario there was no evidence for the numerator mechanism, but there was a tendency for *greater* willingness when more than one exemplar was cued. The direction of this effect is opposite to that predicted by EC theory.

A consideration here, given the differences between scenarios, is why such a scenario effect was not found in Experiments 1 and 2. It is possible that because format was manipulated within subjects in those experiments (rather than the between-subjects manipulation used in Experiment 3) the effect of format was magnified in both scenarios thus masking possible differences between scenarios. That is, participants in Experiments 1 and 2 may have compared the information provided by frequency and probability formats thus increasing the weight placed on judgments driven by frequency information.

*Imaginability ratings.* Figure 4 displays the mean imaginability ratings for the four scenarios (lottery, game show, vaccine and laser, respectively). Taking the two positive scenarios first, the figure suggests a marginal
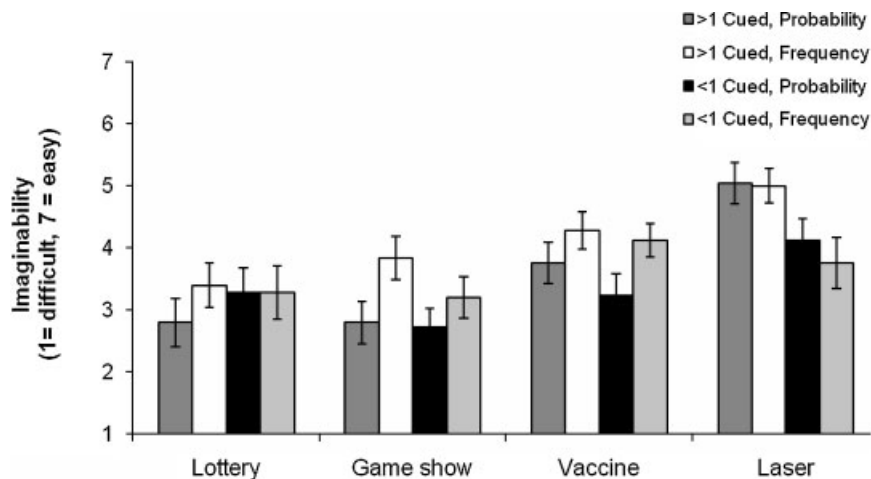


Figure 4. Mean imaginability ratings for each scenario in Experiment 3. Although both factors were manipulated between subjects, bars are presented grouped for each scenario to facilitate cross-scenario comparisons. Error bars are standard error of the mean

(lottery) or more systematic (game show) effect of increased imaginability when a frequency format was used (the average of the lighter bars is higher than the dark bars). There appears to be no effect of number of exemplars cued on imaginability. Statistical analyses confirmed this impression. There were no effects of number of exemplars cued in either scenario ($Fs < 1$), and only in the game show scenario was format significant, $F(1, 96) = 5.28$, $p < .05$ (lottery, $F < 1$). The interaction was not significant in either scenario.

Turning to the negative scenarios, the data for the vaccine scenario suggests that frequency formats do increase imaginability (lighter bars higher than dark) but this effect is not present in the laser scenario. Number of exemplars cued, in contrast, appears to have no effect in the vaccine scenario, but in the laser scenario when more than one exemplar is cued imaginability does appear to be facilitated (bars on the left higher than those on the right). These effects were confirmed by statistical analyses. Format was significant in the vaccine scenario $F(1,96) = 5.01$, $p < .05$, but not in the laser scenario ($F < 1$); number of exemplars cued was significant in the laser scenario $F(1, 96) = 9.85$, $p < .05$, but not in the vaccine scenario ($F < 1$). The interaction was not significant in either scenario.

Taken together, the results from both positive and negative scenarios indicate that, although the effect may not be very strong, there is some support for the claim that frequency formats enhance imaginability. The effect of number of exemplars cued on imaginability is less clear. It has no effect in the positive scenarios but imaginability does increase with the number of exemplars cued in one of the negative scenarios.

*Correlational analysis.* To explore the relation between formats, willingness and imaginability in more detail we conducted correlation analyses. We examined the relation between imaginability and willingness separately for the probability and frequency format versions of the positive and negative scenarios (We collapsed across number of exemplars cued as this variable appeared to have little systematic effect on the two measures.). Table 2 displays the Pearson correlation values for the two measures in the four types of scenario—frequency positive, frequency negative, probability positive and probability negative. Taking the frequency formats first, there is a clear indication that willingness ratings and imaginability ratings are moderately correlated. For the positive scenarios, the higher one's rating of imaginability, the more willing one is to participate; for negative scenarios the reverse is true—the higher the imaginability rating, the lower the willingness.

These patterns of correlations are consistent with claims for the mechanisms underlying the frequency format effect. However, the format account also predicts that these relations should not exist, or at least be weaker, when probability formats are used. Table 2 also displays the correlations between imaginability and willingness for the probability formats. The absence of a significant correlation in the negative scenarios supports the prediction of the format account. However, rather surprisingly, the correlation in the positive scenarios, although weaker than in the frequency format, remains significant in the probability format conditions.

In summary, this analysis provides support for the claim that willingness ratings are correlated with the ease with which a particular scenario can be imagined. In both positive and negative scenarios correlations

Table 2. Pearson correlations between ratings of willingness and imaginability in Experiment 3

| Format and scenario type | Pearson correlation coefficient | *p*-values |
|---|---|---|
| Frequency positive | .436 | .002 |
| Frequency negative | −.404 | .004 |
| Probability positive | .330 | .019 |
| Probability negative | −.196 | .173 |

*Note*: Ratings are collapsed across the two positive and two negative scenarios and the number of exemplars cued factor.

between willingness and imaginability were stronger when scenarios were presented in frequency than in probability formats.

*Regression analysis.* Regression analyses were conducted to examine the relative importance of imaginability and emotion ratings in determining willingness ratings. Although our primary aim for including emotion ratings was to use them as a manipulation check to ensure that the positive and negative scenarios elicited different perceived emotions, by including them in a regression analysis further insight into the mechanisms underlying the format effect might be gained.

In keeping with the correlation analyses, separate regressions were performed for the probability and frequency format versions of the positive and negative scenarios (collapsing across the number of exemplars cued variable). Taking the frequency formats first, in both the positive and negative scenarios the imaginability ratings were significant predictors of willingness (positive: beta $= .418$, $t = 3.23$, $p = .002$; negative: beta $= -.351$, $t = -2.55$, $p = .014$); but the emotion ratings were not (betas $< .19$, $ps > .15$). For the probability formats, in the negative scenarios, neither imaginability nor emotion ratings significantly predicted willingness (imaginability, $p = .17$; emotion, $p = .621$), but for the positive scenarios both imaginability and emotion ratings were significant predictors of willingness (imaginability: beta $= .303$, $t = 2.44$, $p = .018$; emotion: beta $= .412$, $t = 3.31$, $p = .002$).

The regression analyses show that when frequency formats are used ease of imagining is more important than perceived emotion for predicting willingness to participate in events—be they positive or negative. This result is clear-cut, consistent with the prediction of the format account and appears to clarify that imaginability rather than perceived emotion is the principal predictor of willingness ratings. However, the pattern from the probability format scenarios challenges this interpretation. Here the format account would predict that the importance of imaginability and emotion would be reduced relative to the frequency scenarios. This was true for the negative scenarios (in which *neither* were significant predictors of willingness) but not for the positive scenarios (in which *both* were significant predictors). Perhaps the best that can be said is that this pattern of results reflects the difficulty of measuring imaginability and perceived emotion and the idiosyncratic effects of particular fictitious scenarios—issues which we return to in the general discussion.


## GENERAL DISCUSSION

In three experiments, we investigated contrasting accounts of how people make judgments about low-probability events. In the following sections, we discuss the implications of our results for the specific mechanisms of the format account and the EC theory.


### Format effects

In Experiment 1 and in the stronger-positive-emotion scenario of Experiment 3 when the low-probability event was positive (e.g. winning a lottery), participants were significantly more willing to engage in the proposed behaviour when frequency formats were used. In Experiment 2 and in the stronger-negative-emotion scenario of Experiment 3 when the low-probability event was negative (e.g. getting scarred by a laser), participants were significantly less willing to engage in the behaviour when frequency formats were used. These format effects are consistent with similar findings in the literature and are consistent with the notion that frequency formats evoke greater imagery than probability formats (Finucane et al., 2003; Loewenstein et al., 2001; Slovic et al., 2000, 2002; Yamagishi, 1997). A strength of the demonstration of format effects in this paper is that they have been shown in both between and within-subject designs. Some researchers have argued that within-subject factorial designs of the kind used in Experiments 1 and 2 are

inappropriate in many areas of psychological research because the designs provide cues to participants about the intentions of the experimenter (e.g. factors that are varied systematically must be relevant to the dependent variable, Kahneman, 2003). Demonstrating that the format effect holds in both within and between-subject designs is testament to the reality of the effect.

Typically the format effect is discussed with reference to negative emotions with the notion being that judgments about the likelihood of an event occurring are based on 'feelings of risk', with stronger feelings leading to proportionally higher judgments. The proposed mechanism is that the frequency formats, because of the vivid imagery they promote, lead to even stronger feelings of risk than the numerically equivalent, but less concrete, probability representations (Loewenstein et al., 2001; Slovic et al., 2000). Experiment 1 and 3 demonstrated that the format effect works in some positive scenarios as well as negative, presumably due to the stronger 'feelings of benefit' evoked by the frequency representations.

Experiment 3 sought direct evidence that (1) the positive and negative scenarios elicited different emotional reactions and (2) that the different formats affected the ease with which participants could imagine the scenarios. The emotion ratings provided clear evidence that positive scenarios elicited more positive emotions (happy, excited) and negative scenarios elicited more negative emotions (scared, sad). These data also showed that within scenarios, the lottery scenario was perceived as the more positive and the laser scenario was the more negative. Interestingly, these differences in emotional reaction appeared to be a determinant of the reliability of the format effect on willingness ratings.

The imaginability data provided some evidence that scenarios presented in frequency formats were easier to imagine than those presented in probability formats. The relation between imaginability and format was not as clear as some accounts might predict. For one, the relation seemed to be affected by idiosyncrasies of particular scenarios; however, the correlation and regression analyses provided converging evidence that ease of imaginability and willingness were more strongly and consistently related in frequency formats than probability formats.

This is the first direct evidence (that we are aware of) of the proposed mechanism underlying the format effect. As noted earlier, support for the frequency account mechanism is typically indirect. An unpublished study (cited in Slovic et al. (2000)) is an exception to this rule, and it appears to support the results found in the current paper. Participants were given scenarios in which patients were described as having either a '10% probability of committing a violent act' or in frequentist terms as '10 out of 100 similar patients are estimated to commit an act of violence'. They were then asked to 'write a few brief thoughts or images that come to mind as you evaluate the risk posed by this patient' (p. 289). The frequency format produced images of violent acts in participants' reports (e.g. 'Some guy going crazy and killing someone'), whereas the probability format did not.

Other studies which have not manipulated the format of statistics directly, also find some support for the general relation found here between imagery and judgment. For example, Slovic et al. (1991) examined peoples' preferences for vacation locations using a word association technique to elicit imagery. Participants were presented with the name of a place (e.g. San Diego), asked to report the first image or thought that came to mind (e.g. zoo) and then to rate the affective quality of the image. Slovic et al. (1991) found that the summed image ratings were highly predictive of expressed preferences and the location of the vacations subsequently taken by participants.

### Format and emotion

Frameworks such as the affect heuristic (Slovic et al., 2002) propose that, 'availability may work not only through *ease* of recall or imaginability, but because remembered and imagined images come tagged with affect' (p. 414). Such a proposal implies that in addition to enhancing the imaginability of scenarios, frequency formats will also increase the intensity of evoked emotional reactions. Note that this issue is different to the one addressed in the manipulation check analysis. There the interest was simply in whether

positive and negative scenarios elicit different emotions. We found that the scenarios did have the desired effect and that the 'intensity' of the perceived emotion appeared to be a determinant of the reliability of the format effect on willingness. Here, however, the interest is in whether intensity of emotions is affected by the format in which statistical information is presented.

The short answer to this question is 'no'. We did not report a full breakdown of the emotion ratings because they were not informative. The emotion ratings were unaffected by format and the number of exemplars cued (all $F$s < 1). However, it is possible that we were unable to detect any differences due to floor and ceiling effects in the emotion ratings. The positive scenarios elicited a mean rating of 5.62 on the 1–7 scale and the negative scenarios elicited a rating of 1.97. The proximity of these ratings to the end-points of the scale leaves open the strong possibility that potential differences between conditions were masked. The regression analyses also failed to pinpoint a clear role for emotion in predicting willingness ratings.

In summary, we found an asymmetrical relation between emotion, willingness and format. Although the strength of the perceived emotion appeared to be a determinant of the reliability of the format effect on willingness ratings, the format of the information itself (frequency or probability) did not impact on the level of the emotion experienced. Clarifying this relation between emotion, format and decision-making remains a goal for future research. We speculate that a greater range of scenarios and methods for presenting risk information and for eliciting and measuring perceived emotions and ease of imagining will be required to investigate this complex relationship. As we have noted throughout, this paper presents one of the first attempts to measure these constructs in risk-perception experiments. We argue that this represents an advance in helping to clarify the mechanisms underlying the format effect but acknowledge that the 'direct' method of measurement (i.e. simply asking for ratings) may not be the best way to gauge the imaginability and emotional content of fictitious scenarios. These complicated and intriguing questions await the ingenuity of future researchers.

### Evidence for exemplar cuing theory

EC theory proposes two specific mechanisms for enhancing imaginability of events (1) the 'multiplicative'— exemplars are cued only when the product of the reference class and incidence rate is greater than one (regardless of format) and (2) the 'numerator'—integer rather than fractional numerators promote EC.

### Multiplicative mechanism

We found little support for the multiplicative mechanism of EC theory. First, we found clear effects of format, as discussed above, even when the 'problem-relevant sample space' or reference class was held constant across format representations. This effect is in direct contrast to the prediction of the multiplicative mechanism. Second, when an effect of the number of exemplars cued on willingness was found it was in the opposite direction to that predicted by the theory. In Experiment 1, participants were more willing to buy lottery tickets or call a TV station when no exemplars of other 'winners' could be generated than when other winners could be generated (though this effect was not replicated in Experiment 3). In Experiment 3, participants were more willing to participate in a vaccine trial when *more* than one exemplar of other side-effect sufferers was cued. In the remaining conditions, exemplars cued had no significant effect on willingness. The only result that is consistent with the EC interpretation of the EC effect was the enhanced imaginability of the laser scenario when more than one exemplar was cued (Experiment 3).

One potential reason for the discrepancy between our failure to find EC effects like those reported by Koehler and Macchi (2004) is that in our tests of the multiplicative mechanism the product of the reference class and the incidence rate was not sufficiently small. Note that in the Koehler and Macchi study the product was 0.005 compared to those of 0.2 or 0.5 in our scenarios. EC theory does not differentiate between such cases, indeed, as it is formulated an effect of cuing should be observed for a product of 1.0 but not for one of

0.999. Furthermore, the theory predicts the same cuing advantage whether one or 1000 exemplars are cued (Koehler & Macchi, 2004). Notwithstanding these rather counter-intuitive predictions, it is plausible that participants in our experiments rounded up 0.2 and 0.5 to 1, thus facilitating the production of exemplars in both the greater and less than one conditions. This account does not explain the reversal of the exemplar effect found in Experiment 1, but it is possible that the reversal was due to some participants misunderstanding the nature of the lottery and game show scenarios in the <1 conditions. In Experiment 3 when an instruction clarifying that there would not always be a winner was added, the reversed exemplar effect was not found.

One other potential reason for our failure to find evidence for the multiplicative mechanism is that the wording we used in our scenarios did not convey the nature of the 'target' adequately. Our aim was to provide a clear indication of the reference class and the incidence rate in each case, thus facilitating the possible multiplication of the two numbers. A simple interpretation of EC theory is that when the product of this multiplication exceeds 1, exemplars will be generated. However, a slightly more subtle prediction of the theory is that this multiplication will only take place when a multiple target is identified in the wording of the statistic (Koehler & Macchi, 2004).

In our scenarios, arguably, the target was single because we used statements of the form, '50 tickets are sold each day and it is estimated that you have a 1% chance of buying a winning ticket'—although it is clear that the reference class is 50 and the incidence rate is 1% one could claim that focussing on the 1% chance of buying a winning ticket makes the target of the statement single. In contrast if we had used statements of the form, '1% of the 50 tickets sold each day are winning tickets'—the target might have been more readily interpreted as multiple tickets. While we acknowledge this possibility, we do not feel that it undermines the logic of our test of the multiplicative mechanism—namely, testing that when the reference class and incidence rate are presented, exemplars will be generated if their product exceeds one. The results of all three experiments appear to be contrary to this claim.[2]

### Numerator mechanism

The strongest evidence for the numerator mechanism comes from Experiment 2 in which participants were significantly less willing to participate in clinical trials with low-probability negative outcomes (e.g. scarring) when an integer numerator (20 or 50) was used to specify the number of those outcomes than when a fractional (0.2 or 0.5) one was used (The effect was weaker but in the same direction in the laser scenario of Experiment 3.). This numerator effect is consistent with Koehler and Macchi's (2004, Experiment 2) demonstration that DNA match statistics with integer numerators were less convincing of guilt than those with fractional numerators. The EC theory interpretation of this result is that the integer numerator of the incidence rate by itself provides readily available exemplars of co-incidental matches. These other exemplars appear to strengthen the defendant's claim to innocence. In the case of the vaccine/laser treatment, presumably the incidence rate readily provides exemplars of other sufferers of the side effect thus making the treatment less attractive. The imaginability ratings from Experiment 3 provide convergent evidence for this explanation. In the laser scenario, being scarred by the treatment is rated as more imaginable when an integer numerator was used ($M = 5.00$) than when a fractional one was used ($M = 3.76$) (See Figure 4.). This is evidence is, of course, also consistent with the format account.

---

[2]Koehler and Macchi (2004, p.541) state, 'In the lottery example, if a multiple target is provided (e.g. 5000 tickets are sold) and the chance that a particular ticket will win is 0.1%, then the winning-ticket exemplars are likely to be cued (because approximately 500 tickets in this lottery will be winners). However, if the multiple target indicates that only five tickets are sold, then exemplars are not likely to be cued. Algebraically, if the product of the reference-class and the event probability is greater than 1 exemplars are likely to be cued'. We contend that the scenarios we used which clearly identified the number of tickets sold and the event probability satisfy these constraints regardless of any potential ambiguity about the multiple/single nature of the targets.

### Rethinking exemplar cuing

One final point to consider is that Koehler and Macchi's own evidence for EC and the existence of a multiplicative mechanism in particular can be challenged. In their Experiment 1, the key result supporting the multiplicative mechanism was the finding that DNA evidence was less convincing when the perpetrator was thought to reside in a town of 5 000 000 residents (large reference class) than in town with 500 residents (small reference class). The incidence rate in both conditions was held constant (0.001%) and thus although the results are consistent with the notion that participants engage in some multiplication of the incidence rate and the reference class, they are equally consistent with the idea that participants simply anchor on the reference class and reason that if there are only 500 people in the town then there is a lower chance that the police have 'got the wrong man'.

This independent contribution of the reference class could also explain why, in some cases we found a *reversal* of the exemplar cued effect. Displays of greater willingness to participate in a positive event when *fewer* than one exemplar is cued (Experiment 1) and greater willingness to participate in a negative event when *more* than one exemplar is cued (Experiment 3, vaccine scenario) could be explained by participants focusing solely on the size of the reference class. It is plausible that participants estimate the chance that someone else—rather than themselves—might win the lottery or enter the game show and then use this estimate to determine their willingness to participate. Reasoning like this could lead to a decrease in the willingness to participate as the likelihood of generating other 'winners' increases—as we observed. The participants appear to adopt a pessimistic frame of mind, reasoning that: 'if all these other tickets can win then mine won't'. By the same token, in at least one negative scenario (Experiment 3, vaccine), participants appear to adopt an optimistic frame of mind, reasoning that: 'if all these other people suffer side effects then I won't'.

Such an explanation is consistent with the idea proposed by a number of researchers that people focus on the absolute number of ways in which an event can happen rather than the probability ratio associated with an event (e.g. Denes-Raj & Epstein, 1994; Kahneman & Miller, 1986; Kahneman & Tversky, 1982; Miller et al., 1989). One possibility is that in the context of the positive event participants erroneously treat the reference class alone as an indication of how many other possible 'winners' there are and thus indicate higher willingness when that number is low. Similarly in the context of a negative event, the reference class indicates the number of other possible 'losers/sufferers' and thus indicate higher willingness when that number is high. It is important to stress, however, the speculative nature of these explanations given inconsistencies across scenarios and the possibility that confusion about the nature of the positive scenarios in Experiment 1 might have led to the reversed cuing effect.

## CONCLUSION

Our results present some clear challenges to the specific formulation of EC theory provided in Koehler and Macchi (2004). In particular we found little evidence for the multiplicative mechanism that they propose. However, our findings are consistent with the general notion that the imaginability of outcomes influences how people think about low-probability events. In addition, we have provided evidence that this imaginability is enhanced by the use of frequency formats and that imaginability is an important determinant of probability judgments (as indexed by willingness). Future work should aim to clarify further the locus of these effects, their complex relation with emotion, and determine their implications for the important practical problems inherent in risk communication.

## APPENDIX

In Experiments 1 and 2 participants received two questions relating to different scenarios. For all participants one question was in a frequency format and one in a probability format. For half the participants the questions both cued exemplars, for the other half they did not cue exemplars (according to the predictions of EC

theory). In Experiment 3 participants were given four scenarios which were all in a single format and which either did or did not cue exemplars. Below are examples of the scenario questions used.

**Lottery scenario, probability format, <1 exemplar cued**
You are offered the opportunity, today, to buy a ticket in the new 'Sydney Lottery'. Fifty tickets are sold each day and it is estimated you have a 1% chance of buying a winning ticket.

**Game show scenario, frequency format, <1 exemplar cued**
You have the opportunity, today, to enter a TV game show that offers fantastic prizes. Twenty people call the TV station each day to try to get on the show and it is estimated that on the average 0.2 people in every 20 callers are selected.

**Vaccine scenario, probability format, >1 exemplar cued**
Imagine that you are very susceptible to catching flu. A new flu vaccine has been developed which is claimed to guard against all known strains of flu. You are asked to take part in a clinical trial of this new vaccine. Two thousand people worldwide will participate in the trial and you are told that on the basis of previous research with similar vaccines there is an estimated risk of 1% that the vaccine will cause an unpleasant allergic reaction.

**Laser treatment scenario, frequency format, >1 exemplar cued**
Imagine that you have a large birthmark on your face. A new laser treatment for removing birthmarks has been developed and you are offered the chance to participate in a clinical trial of this new treatment. Five thousand people worldwide will participate in the trial and you are told that based on previous research with similar treatments it is estimated that 50 out of 5000 people in a trial will be permanently scarred by the treatment.

REFERENCES

Denes-Raj, V., & Epstein, S. (1994). Conflict between intuitive and rational processing: When people behave against their better judgment. *Journal of Personality and Social Psychology*, *66*, 819–829.
Epstein, S. (1994). Interaction of the cognitive and psychodynamic unconscious. *American Psychologist*, *49*, 709–724.
Finucane, M., Peters, E., & Slovic, P. (2003). The dance of affect and reason. In S. L. Schneider , & J. Shanteau (Eds.), *Emerging perspectives on judgment and decision research* (pp. 327–364). Cambridge: Cambridge University Press.
Finucane, M., Alkhami, A., Slovic, P., & Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making*, *13*, 1–17.
Kahneman, D. (2003). A perspective on judgment and choice. *American Psychologist*, *58*, 697–720.
Kahneman, D., & Tversky, A. (1982). Variants of uncertainty. In D. Kahneman , P. Slovic , & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 509–520). Cambridge: Cambridge University Press.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, *92*, 136–153.

Koehler, J. J. (1997). One in millions, billions and trillions: Lessons from *People v. Collins* (1968) for *People v. Simpson* (1995). *Journal of Legal Education*, *47*, 214–223.

Koehler, J. J. (2001). When are people persuaded by DNA match statistics? *Law and Human Behavior*, *25*, 493–513.

Koehler, J. J., & Macchi, L. (2004). Thinking about low probability events: An exemplar cuing theory. *Psychological Science*, *15*, 540–546.

Lipkus, I., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making*, *21*, 37–44.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, E. S. (2001). Risk as feelings. *Psychological Bulletin*, *127*, 267–286.

Macchi, L. (2000). Partitive formulation of information in probabilistic problems: Beyond heuristics and frequency format explanations. *Organizational Behavior and Human Decision Processes*, *82*, 217–236.

Miller, D. T., Turnbull, W., & McFarland, C. (1989). When a coincidence is suspicious: The role of mental simulation. *Journal of Personality and Social Psychology*, *57*, 581–589.

Purchase, I. F. H., & Slovic, P. (1999). Quantitative risk assessment breeds fear. *Human and Ecological Risk Assessment*, *5*, 445–453.

Slovic, P., Finucane, M., Peters, E., & MacGregor, D. G. (2002). The affect heuristic. In T. Gilovich , D. Griffin , & D. Kahneman  (Eds.), *Heuristics and biases* (pp. 397–420). New York: Cambridge University Press.

Slovic, P., Layman, M., Kraus, N., Flynn, J., Chalmers, J., & Gesell, G. (1991). Perceived risk, stigma, and potential economic impacts of a high-level nuclear waste repository in Nevada. *Risk Analysis*, *11*, 683–696.

Slovic, P., Monahan, J., & MacGregor, D. G. (2000). Violence risk assessment and communication: The effects of using actual cases, providing instruction and employing probability versus frequency formats. *Law and Human Behavior*, *24*, 271–296.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*, 207–232.

Yamagishi, K. (1997). When a 12.86% mortality is more dangerous than a 24.14%: Implications for risk communication. *Applied Cognitive Psychology*, *11*, 495–506.

*Authors' biographies*:

**Ben Newell** is a Senior Lecturer at the University of New South Wales. His research interests include judgment and decision-making, with particular interest in the cognitive mechanisms underlying multiple cue judgment, and the implicit/ explicit distinction in learning, memory, categorization and decision making.

**Chris Mitchell** is a Senior Lecturer at the University of New South Wales. His interests are in human learning, memory and attention. His present focus is the application of associative learning theory to perceptual learning and recognition memory.

**Brett Hayes** is an Associate Professor at the University of New South Wales. His interests include categorization, concept formation, and inductive reasoning in adults and children. His current work focuses on how category representations are revised in the light of new data, and on the cognitive mechanisms that underlie inductive inference.

*Authors' addresses*:

**Ben R. Newell, Chris J. Mitchell** and **Brett Hayes**, School of Psychology, University of New South Wales, Sydney 2052, Australia.