

Imagining Low Probability Events: Contrasting Exemplar Cuing and Frequency Format Accounts

Ben R. Newell (ben.newell@unsw.edu.au)

School of Psychology, University of New South Wales, Sydney, Australia

Chris Mitchell (chris.mitchell@unsw.edu.au)

School of Psychology, University of New South Wales, Sydney, Australia

Brett K. Hayes (b.hayes@unsw.edu.au)

School of Psychology, University of New South Wales, Sydney, Australia

Abstract

The extent to which a low probability event can be imagined appears to increase the weight attached to the possibility of that event occurring. Two experiments tested contrasting accounts of how this ‘imagability’ of events is enhanced. The experiments used negative (e.g. suffering the side effect of a vaccine) and positive (e.g. winning a lottery) low probability events. Both experiments found strong support for the frequency format account, whereby imagability is enhanced through the use of frequency formats for conveying statistical information (e.g., 20 out of 2000). However, only limited support was found for ‘exemplar-cuing theory’ (J.J. Koehler & L. Macchi, 2004), which proposes two distinct mechanisms for the generation of instances. Overall, the results support the claim that the imagability of outcomes plays a role in thinking about low probability events, but question the underlying mechanisms specified by exemplar cuing theory for mediating such effects.

Introduction

A fatal shark attack on a 17-year-old boy off West Beach in Adelaide left the popular beaches of South Australia almost deserted this summer. 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 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). In the case of the shark attack, it is not difficult to understand why the photographs, headlines, and TV pictures led to the evocation of terrifying images in the public mind. However, in other less dramatic cases, we can ask under what circumstances these instances of low probability events are generated.

One property that seems to affect the availability of an instance is the ease with which it can be imagined. Instances that can be easily imagined appear to have a stronger influence on judgments than those that cannot be easily imagined (e.g., Epstein, 1994). One factor that is claimed to affect this ‘imagability’ of instances is whether information is presented in terms of frequencies or in terms of

probabilities. 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%). The explanation was that only the frequency presentation generated a ‘terrifying image’ of the recidivist in the mind of the clinician (Slovic et al., 2000). Purchase and Slovic (1999) have documented similar frequency format effects in public assessments of the risks associated with chemical spills.

The imagability idea is also the basis for Koehler and Macchi’s (2004) recent “exemplar cuing theory” (EC). EC states, “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). According to EC, however, the use of a frequency format is not the crucial factor underlying the imagability of exemplars. Koehler and Macchi discuss two separate mechanisms for the facilitation of exemplar generation. The first 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 (e.g., an incidence rate of 1% and a reference class of 500,000 generates 5000 exemplars), but not when it is less than 1 (e.g., 1% and 50 generate 0.5 of an exemplar). Importantly, this mechanism is unaffected by whether incidence rates are provided as percentages (1%) or frequencies (1 out of 100) (see Macchi, 2000 for more discussion of this claim). The second mechanism cues exemplars when the numerator of the incidence rate is 1 or more (e.g., 20 out of 2000) but not when it is less than 1 (e.g., 0.2 out of 20) (cf. Denes-Raj & Epstein, 1994; Epstein, 1994).

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, other than the defendant, 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 also 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 1 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 were cued. Importantly, the frequency/probability manipulation did not have a significant effect on judgments.

In the second experiment 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 juror's attention to other individuals who may match, despite the incidence rates being mathematically identical. Again, this prediction was supported: jurors were relatively less impressed by the evidence when a fractional numerator was used.

Thus the current situation is that there are two contrasting accounts of how the imagability of low probability events is enhanced. One account suggests that the use of frequency formats is crucial to enhancing imagability (e.g., Purchase & Slovic, 1999; Slovic et al. 2000); the other proposes two specific mechanisms – the 'multiplicative' and the 'numerator' – that increase the likelihood that images are cued, independent of the format employed (Koehler & Macchi, 2004). In the two experiments we report we created a design in which we could test 1) the contrasting predictions of the frequency and exemplar cuing accounts regarding the format effect, and 2) the specific mechanisms proposed by the exemplar cuing account.

The design was as follows: Question format (probability or frequency) was manipulated within subjects, and Number of exemplars cued (less than 1 or greater than 1) was manipulated between subjects. The multiplicative mechanism of the exemplar cuing account was tested by comparing responses to probability format questions in which the product of the reference class and the incidence rate was greater than one (e.g., 2000 and 1%), with questions in which the product was less 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 1 (e.g., 50) to those in which it was less than 1 (e.g., 0.5).

In this design, the exemplar cuing 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 a number of exemplars cued effect, 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) (c.f., Epstein, 1994; Slovic et al., 2000).

An additional feature of these experiments is that they test the generality of EC theory. Although they only tested the theory in the context of DNA evidence, Koehler and Macchi (2004) speculated that the effects they observed 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, we expect to see exemplar-cuing effects in such scenarios that are consonant with those observed previously.

Experiment 1: Negative Events

Experiment 1 tested the predictions of the two accounts 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. The frequency account predicts that willingness should only decrease when a frequency format (which enhances the imagability of other sufferers) is used.

We manipulated the number of exemplars cued (according to EC) by a) varying the number of individuals involved in the trials (i.e., problem space or 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 experiment was a 2 (Number of Exemplars Cued: >1, <1) x 2 (Format: probability, frequency) mixed design with repeated measures on the second factor. Participants were provided with a questionnaire containing the two questions about willingness to participate in clinical trials of a flu vaccine and a laser treatment for birthmark removal. 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 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.

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). The incidence rate of the negative event (vaccine: unpleasant allergic reaction; laser: permanent scarring) was given as a probability (1%) or as the equivalent raw frequency (e.g. 20 out of 2000 or 0.2 out of 20). This resulted in a question in which the product of the reference class and the incidence rate was greater than one (e.g. 2000 and 1%), a question in which the product was less than one (e.g. 20 and 1%), a question in which the numerator cued exemplars (e.g. 20 out of 2000) and a question in which it did not (e.g. 0.2 out of 20). Table 1 displays the design of Experiment 1 with the figures used in the vaccine scenario given as examples of the question types. The Appendix provides examples of the questions used.

Table 1: Design of Experiment 1

No of Exemplars Cued	Question Format	
	Probability	Frequency
> 1	1%, 2000 (20)	20, 2000 (20)
< 1	1%, 20 (0)	0.2, 20 (0)

Note: Number of exemplars cued (according to EC Theory) for each question type is given in parentheses

After reading the scenario participants made a “willingness to participate in the trial” 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 the pattern of effects of the exemplar cuing and format factors was very similar for each of the scenarios and so the data were collapsed across the two in all subsequent analyses. Figure 1 displays the mean willingness ratings collapsed across the flu vaccine and laser treatment scenarios. A 2 (Number of exemplars cued: >1, <1) x 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 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 only had an effect on willingness when the frequency format was used, $F(1, 139) = 4.19, p < .05$. (Probability format $F < 1$). That is, when a frequency format was used, participants were significantly less willing to participate in the clinical trials when an integer numerator was used than when a fractional one was used.

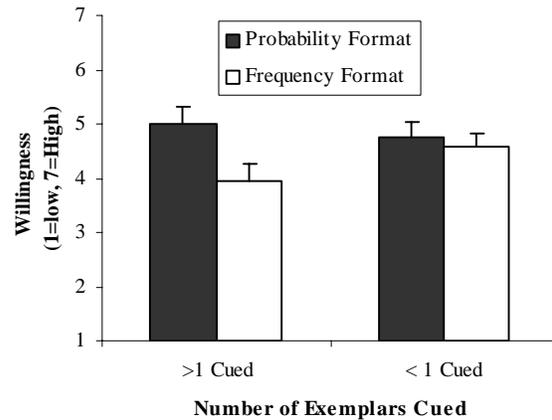


Figure 1. Mean willingness ratings collapsed across the two negative scenarios in Experiment 1

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. The interaction of Question format with Number of exemplars cued, along with the main effect showing lower willingness overall in frequency format conditions cannot be accounted for by EC theory (Koehler & Macchi, 2004; Macchi, 2000) but supports the ‘vivid imagery’ explanation of the probability/frequency format distinction offered by Slovic et al., (2000). That is, frequency formats, which facilitate imagining other individuals who have been scarred or become ill (e.g. 20 out of 2000), reduce the willingness to participate in the clinical trials.

Experiment 2: Positive Events

Experiment 2 tested the predictions of the two accounts in scenarios in which the low probability event was a positive outcome for the participant. Participants were asked to indicate their willingness to participate in a lottery and to call a TV station for the chance of participating in a game show. In the case of the lottery the low probability event was winning, 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 of the two proposed mechanisms 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 facilitate imagability) are used.

Method

Participants

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

Design and Procedure

The design of Experiment 2 was identical to Experiment 1 with the exception that participants were given positive event scenarios: buying a lottery ticket and entering a TV game show.

The reference class sizes, incidence rates and numerators were manipulated in the same way as in Experiment 1. All numerical values were identical. The 0.2, 20 and 2000 values used for the TV game show and the 0.5, 50 and 5000 for the lottery ticket. 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. Examples of the questions used are given in the Appendix. Participants used a 7-point Likert scale to rate their willingness to buy a ticket in the lottery and willingness to call the TV station to enter the game show.

Results and Discussion

Consistent with Experiment 1, preliminary analysis indicated a similar pattern of effects in both scenarios and so the data were collapsed across the two in all subsequent analyses. Figure 2 displays the mean willingness ratings collapsed across the lottery and game show scenarios.

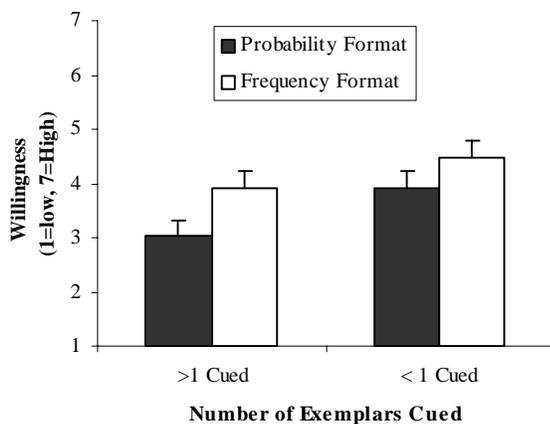


Figure 2. Mean willingness ratings collapsed across the two positive scenarios in Experiment 2

A 2 (Number of exemplars cued: >1, <1) x 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 the theory. That is, participants were *less* willing to engage in the behavior (buy a lottery ticket or participate in a game show) when (according to the theory) more 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 behavior when frequency formats were used. The interaction was not significant ($F < 1$).

The results of Experiment 2 provide no support for EC theory. When, according to the theory, exemplars of other winning tickets or successful game show entrants were cued, participants were significantly less willing to buy a lottery ticket or call the TV station. However, the higher willingness ratings in the frequency format conditions strongly supported the idea that frequency formats facilitate the evocation of other examples of the low probability events (Slovic et al., 2000).

General Discussion

Two experiments investigated contrasting accounts of how the imagability of low probability events is enhanced. EC theory proposes two specific mechanisms for enhancing imagability 1) the 'multiplicative' – exemplars are cued only when product of the reference class and incidence rate is greater than 1 (regardless of format), and 2) the 'numerator' - integer rather than fractional numerators promote exemplar cuing. The frequency account simply suggests that frequency formats lead to more vivid imagery than do probability formats.

We found no evidence to support the multiplicative mechanism of EC theory. In Experiment 1 there was no difference in the willingness to participate in a clinical trial between a condition in which the product was more than 1 (20 or 50) and a condition in which it was less than 1 (0.2 or 0.5). In Experiment 2 the opposite of the predicted pattern was found: participants were more willing when no exemplars of other 'winners' could be generated than when other winners could be generated.

We found limited support for the 'numerator' mechanism. In Experiment 1 participants were significantly less willing to participate in clinical trials with low probability negative outcomes 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. This pattern was not found in Experiment 2: participants were more willing to engage in behaviors with low probability positive outcomes when a fractional numerator was used.

In contrast with the failure to find clear support for EC, results strongly supported the notion that frequency formats evoke greater imagery and affect than probabilistic formats. In Experiment 1 when the low probability event was

negative, participants were less willing to engage in the behavior when frequency formats were used. In Experiment 2 when the low probability event was positive, participants were more willing to engage in the proposed behavior when frequency formats were used. The different findings with the probability and frequency formats lend support to the claims of Kahneman and Tversky (1982), Gigerenzer and Hoffrage (1995), Slovic et al. (2000) and others that thinking about frequencies differs from thinking about probabilities. (See Sloman, Over, Slovic and Stibel 2003 for an illuminating discussion of the basis for this difference).

One possible reason for the discrepancies between our results and those of Koehler and Macchi 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. It is plausible that participants rounded up these larger figures to 1, thus facilitating the production of exemplars in both conditions. Such an account would explain why we found no difference between the probability format conditions of Experiment 1, but it does not explain the reversal of the exemplar effect in the probability conditions of Experiment 2.

This latter reversal could be explained within the EC framework by suggesting that participants could be first estimating 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 reversal effect was also present in the frequency format conditions of Experiment 2. Although there was a main effect of frequency over probability, participants were more willing to participate in positive events when a fractional numerator was used than when an integer was used. One explanation for the reversal in both conditions is that the reference class is an independent contributor to the willingness to engage in low probability positive events. This explanation seems a little ad hoc, but is based on the very intuitive idea that people are simply more willing to play in lotteries when the number of other tickets sold is small. This would explain the observed pattern of greater willingness when the lottery was described in terms of 50 as opposed to 5000 other tickets being sold in a day. This factor appeared to increase willingness over and above the effect of presenting the data in a frequency format.

These alternative explanations highlight the need to document the format effect in more detail. It seems that it is not only the use of frequency formats per se that is important, but also the aspects of the information that people focus on, and how imagined outcomes are subsequently used in reasoning about particular scenarios. Further experimentation is required to flesh out these alternative accounts.

In conclusion, our results present some clear challenges to the specific formulation of EC theory provided in Koehler and Macchi (2004). In particular we found no evidence for the multiplicative mechanism that they propose. However, our findings are consistent with the general notion that the imagability of outcomes affects how people think about low probability events and that this imagability is enhanced by the use of frequency formats. Further work is required to define the locus of these effects and determine their implications for the important practical problems inherent in risk communication.

Acknowledgments

This research was supported a Faculty Research Grant from the University of New South Wales.

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.
- Gigerenzer, G. & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: frequency formats. *Psychological Review*, 102, 684-704.
- 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
- 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., & Macchi, L. (2004). Thinking about low probability events: an exemplar cuing theory. *Psychological Science*, 15, 540-546.
- 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.
- Purchase, I.F.H. & Slovic, P. (1999). Quantitative risk assessment breeds fear. *Human and ecological risk assessment*, 5, 445-453.
- Sloman, S.A., Over, D., Slovic, L., & Stibel, J.M. (2003). Frequency illusion and other fallacies. *Organizational Behavior and Human Decision Processes*, 91, 296-309.
- 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.

Appendix

In both experiments 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. Below are example questions for both experiments. For Experiment 1 examples of the questions that cued exemplars are given, for Experiment 2 the 'non-cuing' questions are provided.

Experiment 1: *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. 2000 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.

Experiment 1: *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. 5000 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.

Experiment 2: *Lottery Scenario, Probability Format, < 1 exemplar cued.*

You are offered the opportunity, today, to buy a ticket in the new "Sydney Lottery". 50 tickets are sold each day and it is estimated you have a 1% chance of buying a winning ticket.

Experiment 2: *Game Show Scenario, Frequency Format, < 1 exemplar cued.*

You have the opportunity, today, to enter a TV game show that offers fantastic prizes. 20 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.