Title
Comparing Risk Reductions: On the Interplay of Cognitive Strategies, Numeracy, Complexity, and Format.

Authors
Barton, Adrien
Cokely, Edward
Galesci, Mirta
et al.

Peer reviewed

Comparing Risk Reductions:
On the Interplay of Cognitive Strategies, Numeracy, Complexity, and Format

Adrien Barton (barton@mpib-berlin.mpg.de)
Edward T. Cokely (cokely@mpib-berlin.mpg.de)
Mirta Galesic (galesic@mpib-berlin.mpg.de)
Anna Koehler (koehler@mpib-berlin.mpg.de)
Mario Haas (haas@mpib-berlin.mpg.de)

ABC, Max Planck Institute for Human Development,
Lentzeallee 94, 14195 Berlin, Germany

Abstract

Effective communication of risks requires a theoretical understanding of the influences of cognitive and task constraints. We present progress toward the identification of performance parameters in an experiment that explores the dynamic interplay of four key risk communication variables (i.e., cognitive strategies, numeracy, complexity, and format). Specifically, we conducted a protocol analysis to trace the types of cognitive strategies used when comparing the helpfulness of two treatments. Variability in cognitive strategies was also examined as related to (1) format (expressed either as absolute or relative risk reductions in either a frequency or single-event probability format); (2) task difficulty; and (3) numerical skill. Results indicated that highly numerate people often effectively used more complex strategies. However, the performance advantage of highly numerate people only existed when comparing two relative risk reductions (which requires a complex strategy), but not when comparing two absolute risk reductions (which requires a simple strategy). A frequency format was also found to produce additional benefits in very difficult tasks (i.e., when comparing absolute and relative risk reductions). Generally, although strategies and accuracy are influenced by many factors, risk communication tended to be most transparent when presented in terms of absolute risk reductions.

Keywords: Risk communication; framing; risk reduction; frequency format; single-event probability; individual differences; numeracy; cognitive strategies; protocol analysis.

Introduction

Absolute and relative risk reductions

How can we effectively communicate risks? The answer depends on the dynamic influences of skills, strategies, and task complexity and formats. For example, many studies have shown that the framing of risk information (i.e. the way the information is formatted) can have a large influence on judgments and decisions (Edwards, Elwyn, Covey, Matthews & Pill, 2001; Gigerenzer, 2003; Tversky & Kahneman, 1981). In particular, two indicators can be used in order to convey risk reductions afforded by a treatment including (1) the absolute risk reduction and (2) the relative risk reduction. Consider two groups of people: One is a trial group taking a treatment and the other is a control group. A definition of risk reduction requires the identification of the base-line risk (BL – i.e., the proportion of participants in the control group who have the disease) and the proportion of participants in the trial group who have the disease (T). Accordingly, an absolute risk reduction is defined as BL-T, and a relative risk reduction as (BL-T)/BL (note that a relative risk reduction is normalized to the base-line risk).

To illustrate potential sources of confusion, a relative risk reduction of 25% could mean a reduction in disease from 40% to 30% (corresponding to an absolute risk reduction of 10%) or from 0.4% to 0.3% (corresponding to an absolute risk reduction of only 0.1%). Relative risk reductions are often larger percentages than absolute risk reductions and thus may be perceived as more effective. For example, a relative risk reduction of 25% may be perceived as more effective than an absolute risk reduction of 0.1%, even when they are equivalent. This can be particularly problematic when the nature of the risk reduction (absolute or relative) is not explicitly stated. Of course, most treatments have negative side effects, which may affect subjective preferences for treatment. Nevertheless, given that relative risk reductions are normalized to the base-line risks they do not enable an accurate estimate of the trade-offs between benefits and side-effects (or costs).

Risk communication and risk reductions are not limited to medical treatments. Generally, the effect of every action which reduces any risk can be expressed as a relative or an absolute risk reduction (e.g., an education program to reduce risky driving behavior). Therefore, it is crucial to develop a detailed theory that will allow effective communication of risks to the public, to policy makers, and to those who routinely deal with risk (e.g., physicians, bankers).

Study of biases in comparing risk reductions

Several studies have shown that physicians and patients favor a treatment, or consider it more helpful, if its beneficial effects are expressed as a relative risk reduction rather than an absolute risk reduction (for a meta-analysis of the scientific literature see Covey, 2007). This could be considered a type of bias. One could however have both normative and descriptive concerns about this claim.

Normative aspects In the meta-analysis of Covey (2007), only three studies out of twenty-eight provided the base-line

risk in addition to the relative risk reduction. Without base-line risk there is no normative criterion enabling the determination of whether participants exhibit bias when comparing a relative with an absolute risk reduction.

When base-line risk is provided, it can be argued that participants are biased if they overestimate the same risk reduction in a relative format than in an absolute format. However, the meta-analysis on these three studies did not provide a clear indication that people favor risk reductions framed in a relative format when base-line risks were provided. Moreover, one cannot determine the direction of the bias. Do people overestimate the efficiency of treatments whose risk reduction is communicated in a relative format? Alternatively, do people underestimate the efficiency when it is communicated in an absolute format? Interestingly, Covey (2007) also questioned whether the relative nature of the risk reduction was sufficiently explicit: In some studies, participants may have interpreted the relative risk reduction as an absolute one. Covey concluded that this bias may be partially attributed to the methodological procedures.

Descriptive aspects To the best of our knowledge, research has yet to directly investigate the cognitive strategies people use when trying to estimate the helpfulness of a treatment on the basis of its risk reduction – whether absolute or relative. The examination of cognitive strategies is a necessary part of ongoing psychological research, in order to assess the most common misconceptions. Moreover, several studies have mixed probability/frequency formats, which can complicate interpretation (i.e., probability on one side and frequency on the other side). For example, Sheridan, Pignone and Lewis (2003), expressed relative risk reduction as probability and absolute risk reductions as a mixture of probability and frequency (for further discussion see Gigerenzer, 2003). If one wants to study the difference caused by relative versus absolute format, one should express all risk reductions consistently (frequency or single-event probability).

Experiment

We conducted an experiment and protocol analysis in order to examine the cognitive strategies people use when comparing different risk reductions in conjunction with the influence of numeracy, format (absolute or relative risk reduction, frequency or single-event probability), and task difficulty. All materials provided base-line risks, and made clear for each risk reduction whether it was absolute or relative, specifying the reference class to which the risk reduction applied.

Method, material and procedures

Fifty-five participants (28 men, 27 women) were recruited at the Max Planck Institute for Human Development (Berlin) and paid 17€ for their participation. Participants completed a demographics questionnaire (e.g., sex, age) and the numeracy test of Lipkus, Samsa and Rimer (2001), which assesses one's general facility with probabilities (note: an

additional item was taken from the scale of Schwartz, Woloshin, Black and Welch, 1997 – the item involving a coin toss).

Structure of each task Every participant was given four risk reduction tasks, each quantifying the helpfulness of two treatments either in a relative or in an absolute risk reduction format. In every task, two diseases were presented. For each of them, the following information was given: the base-line risk (BL), the amplitude of the risk reduction (RR) and the reference class (RC) of the risk reduction. The reference class was presented as either (1) the whole population, which corresponds to an absolute risk reduction; or (2) the population who would otherwise get the disease, which corresponds to a relative risk reduction.

Absolute risk reduction and relative risk reduction The four tasks had the same structure, including independent variables BL_1, RR_1, BL_2, RR_2 (integers) and RC_1, RC_2 (expressions); the instructions read as follows:[1]

"[First paragraph] Consider two different diseases D_1 and D_2 with similar symptoms; for each of them, a particular treatment (which can be for example a vaccine, a medical prescription, a particular lifestyle…) could reduce the risk of contracting the disease. [Second paragraph] BL_1% of the whole population normally get the disease D_1 in their life. But following the treatment A would prevent RR_1% of RC_1. [Third paragraph] BL_2% of the whole population normally contract the disease D_2 in their life. But following the treatment B would prevent RR_2% of RC_2."

The expressions RC_1 and RC_2 read either as "the whole population", which will be abbreviated "Abs" (as this corresponds to an absolute risk reduction); or as "the population who would otherwise get this disease", which will be abbreviated "Rel" (as this corresponds to a relative risk reduction).

Frequency or single-event probability Half of the participants (*N*=27: 15 men, 12 women) received their instruction in a modified single-event probability format; the second and third paragraph then read as follows:

"[Second paragraph] An average person in the whole population has a BL_1% probability of getting the disease D_1 in her life. But following the treatment A would reduce the probability of getting this disease of RR_1% for RC_1. [Third paragraph] An average person in the whole population has a BL_2% probability of getting the disease D_2 in her life. But following the treatment B would reduce the probability of getting this disease of RR_2% for RC_2."

In this single-event probability format, the expressions RC_1 and RC_2 read either as "an average person in the whole population", abbreviated as "Abs"; or as "an average person in the population who would otherwise get the disease", abbreviated as "Rel".

---

[1] The original text was presented in German.

Tasks The following values were used across the four tasks:

Table 1: Independent variables for the four tasks[2]

	BL_1	RR_1	RC_1	BL_2	RR_2	RC_2
T_1	15	4	Abs	20	4	Abs
T_2	30	1	Rel	2	10	Rel
T_3	15	2	Abs	10	1	Rel
T_4	3	1	Abs	5	1	Rel

The order was randomized in the following way: half of the participants received the tasks in the order T_1, T_4, T_3, T_2; and the other half in the order T_3, T_2, T_1, T_4.

Dependent variables At the end of each task, the participants were asked to indicate their answers to the two following questions. *Question 1*. "Without treatment, which disease is more dangerous for the whole population? (1) The disease D_1 is more dangerous than the disease D_2. (2) The disease D_2 is more dangerous than the disease D_1. (3) Both are equally dangerous." *Question 2*. "Which treatment would be more helpful for the whole population? (1) The treatment A would be more helpful for prevention of disease D_1 than the treatment B for prevention of disease D_2. (2) The treatment B would be more helpful for prevention of disease D_2 than the treatment A for prevention of disease D_1. (3) Both treatments would be equally helpful."

Normative strategies The normative strategy for question 1 consists of selecting the disease whose base-line risk BL is higher. Question 1 was only given to prevent an invalid conversational implicature. The maxim of quantity of Grice (1989) enjoins not to make a contribution more informative than is required. If we had only asked question 2, this could have suggested that base-line risks had to be used to answer this question. The addition of question 1 limits the reasons to believe this. This was necessary as a critical issue of our study involve analyses of the answers to question 2.

Question 2 asked which treatment is the most helpful for the *whole population*. Therefore, it required the selection of the treatment with the higher absolute risk reduction. When the effect of the treatment was expressed as a relative risk reduction, one had to multiply this value with the base-line risk in order to get the absolute risk reduction. When both risk reductions were absolute, the normative strategy was to select the treatment which had the highest RR value (this is a simple task). When both risk reductions were relative, the normative strategy was to select the treatment which has the highest $BL \times RR$ value (this is a moderately complex task). And when one risk reduction was absolute and the other relative, the normative strategy was to compare the RR value of the absolute risk reduction with the $BL \times RR$ value of the relative risk reduction, and to select the treatment

associated with the higher of these two values (this is a complex task).

Of note, there is an alternative interpretation of some instructions. Consider for example the task T_3, in which treatment A would prevent 2% of the whole population from getting this disease, whereas treatment B would prevent 1% of the population who would otherwise get this disease from getting it. Participants could think that the 2% of the whole population immunized by treatment A could be people who would anyway not get the disease; whereas treatment B is sure to save at least some people who would otherwise get this disease. This would lead to the selection of the treatment corresponding to a relative risk reduction as the best treatment, because it is *sure* to save some people (classified thereafter as strategy $S5_b$). Although this was not our initial understanding of the text, we acknowledge that it is a possible understanding and a "rational" interpretation.

Classification of strategies To trace cognitive strategies we conducted a concurrent protocol analysis. Participants were instructed to "think aloud" while completing risk comparison tasks. These verbal reports were then identified and coded based on the presence of unique processing products (for an example see Cokely & Kelley, 2009; but for a detailed review see Ericsson & Simon, 1993). Specifically, strategies were coded with respect to their complexity, which was operationalized as follows. In order to compare helpfulness of a treatment, participants could consider two numerical cues including (1) the risk reduction amplitude (RR), and (2) the base-line risk (BL). If a strategy used no more than one cue, it was classified as simple. If a strategy used both numerical cues or took into consideration the reference class RC (whose relevance to the problem is not obvious at a first sight) it was classified as complex. We divided the possible strategies as described in Table 2.

Table 2: Strategies: description and complexity

S1	Select treatment with higher RR value.	Simple
S2	Ethical or intuitive considerations.	Simple
S3	Simple alternative strategies.	Simple
S4	Select treatment with higher $BL \times RR$ value.	Complex
S5	In case one RC is Abs and the other is Rel, either: - $S5_a$: select the treatment which has the highest value of RR (for RC=Abs) or $BL \times RR$ (for RC=Rel). - $S5_b$: select the treatment corresponding to RC=Rel.	Complex
S6	Complex alternative strategies.	Complex
S7	Impossible to identify.	/

The normative strategy for task T_1 is S1; the normative strategy for task T_2 is S4; and the normative strategy for tasks T_3 and T_4 is either S5a or S5b, depending on the

---

[2] Some lifestyle interventions could lead to very low relative risk reduction such as 1% relative risk reduction.

interpretation of the task (it should be noted that both strategies S5a and S5b imply that the participant noticed the difference of reference class implied by the two treatments). Moreover, the strategies S1, S2 and S3 were classified as simple strategies; and the strategies S4, S5 and S6 were classified as complex strategies.

Hypotheses

We hypothesized that people should perform better on tasks that required simpler strategies (e.g., the more they would use the normative strategy); people would then perform better on T_1 than on T_2, and better on T_2 than on T_3 and T_4. We also hypothesized that numeracy would predict complexity of strategies (for similar effects in risky choice see Cokely & Kelley, 2009) and that numeracy would also predict use of normative strategies.

Research has shown that some frequencies can be beneficial in some tasks (natural frequencies for Bayesian tasks, cf. Gigerenzer & Hoffrage, 1995; simple frequencies for conjunction effect, cf. Tversky & Kahneman, 1983). It is currently unclear whether frequencies that are strongly isomorphic to single-event probabilities can improve performance. We therefore also examined whether a frequency format would increase normative strategies (and for whom).

Results

Numeracy Numeracy (0-12) was assessed ($M = 10.1$, $SD = 1.86$). Median split was performed to divide highly numerate ($M = 11.5$, $N = 29$) from less numerate ($M = 8.5$, $N = 26$) participants.

Protocol analysis Verbal reports of answers to question 2 were analyzed independently by two raters (MH and AB) and each classified in one of the seven categories. Across the coding of the 220 strategies (4×55) the inter-rater reliability was substantial ($kappa = 0.61$, $p = 0.001$).[3]

Participants used many different variants of similar strategies. For example, strategies which were coded as S3 included: select the treatment associated with the lowest RR; select the treatment associated with the highest BL. Strategies coded as S6 included: select the treatment associated with the highest ratio RR/BL; select the treatment associated with the highest ratio BL/RR; select the treatment associated with the lowest difference BL-RR. Strategies coded as S2 included: select both treatments, for ethical reasons; choose by gut feelings; choose at random. In S7, we categorized non verbalized strategies (e.g., "I think treatment A would be more helpful" or "[reading the instructions] Treatment B would be more helpful; yes, correct") and ambiguous strategies (e.g., "Most efficient for the whole population? Hmm, 1%, 1% [stammering] because… Prevention, more efficient for the prevention… [stammering]").

Figure 1 displays the proportions of strategies that were classified into the seven categories for respectively tasks T_1, T_2, T_3 and T_4, separated for higher and lower numerate participants.

Numeracy and complex strategy For each participant, we computed, among the answers that were recognized (i.e. categories S1 to S6), the proportion of complex strategies. Highly numerate participants strongly favored more complex strategies ($M = 0.57$, $SD = 0.44$) as compared to low numeracy participants ($M = 0.23$, $SD = 0.27$), $t(41.8) = -3.39$, $p = 0.001$, one-tailed[4], Cohen's $d = 1.1$.
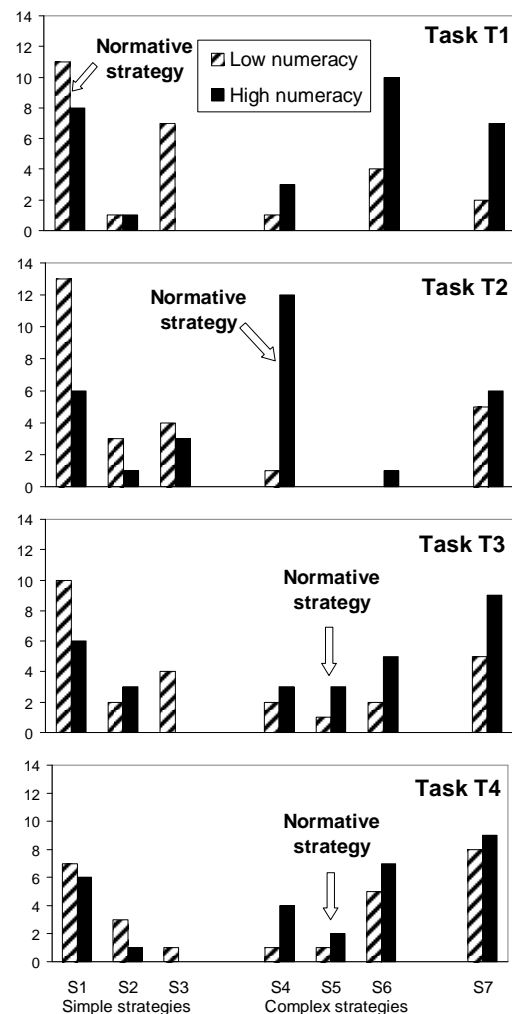

Figure 1: Frequency of strategies used in the four tasks

Numeracy and normative strategy For each participant, we computed the proportion of normative strategies (S1 for T_1, S4 for T_2, and S5 for T_3 and T_4) and non-normative strategies (including non-verbalized and ambiguous strategies). Overall, as the task complexity increased, the

---

[3] Results were similar and significant for either rater. The current data analysis is based on the coding of MH.

[4] Results show correction for unequal variances revealed by Levene's test.

proportion of normative strategies decreased linearly ($F (1,55) = 30.4$, $p < 0.001$, $R^2 = 0.36$).

Figure 2 shows the dependence of the performance on complexity of the tasks and numeracy. Participants low in numeracy only performed well in the simpler task, T_1 (which simply required that one compares risk reduction numbers). When normative strategies were complex, less numerate participants were unable to solve the corresponding task. Highly numerate participants performed well in both simple tasks (T_1) and moderately difficult (T_2) tasks. For example, on the task T_2, highly numerate participants more often used normative strategies ($M = 0.52$, $SD = 0.51$) as compared to participants low in numeracy ($M = 0.05$, $SD = 0.22$), $t (30.3) = -4.06$, $p < 0.001$, one-tailed[4], $d = 1.5$. Both participants high and low in numeracy performed poorly in the most difficult tasks (T_3 and T_4) (see Figure 2).
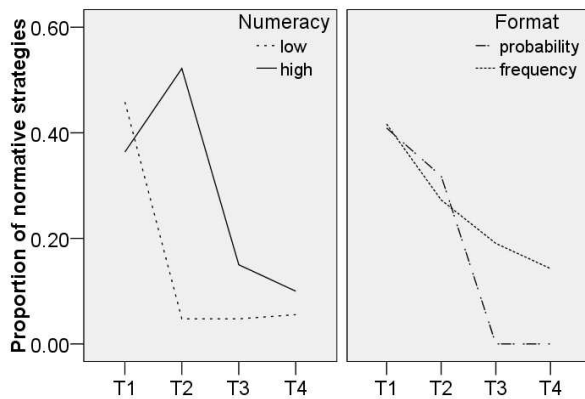

Figure 2: Proportions of normative strategies, depending on numeracy and format

Format and normative strategy Frequency formats did not improve performance on tasks T_1 and T_2, which used the same reference class (absolute or relative) for both risk reductions. However, frequency formats did improve judgment for tasks T_3 and T_4, which mixed absolute and relative risk reductions. Participants in the probability group never answered a single question correctly (Figure 2). When considering T_3 and T_4, a two-way mixed ANOVA showed a moderate-sized performance difference between participants in the frequency and in the single-event probability condition, $F (1,53) = 4.05$, $p = 0.05$, $R^2 = 0.07$.

Discussion

Our results illustrate that the efficacy of risk communication will predictably vary as a function of skill, format, and task difficulty. Our results suggest three major implications: (1) people low in numeracy tend to neglect reference classes; (2) highly numerate people tend to use more complex (but not necessarily normative) strategies in comparing risk reductions, and (3) frequency formats can increase the

likelihood of attending to differences between reference classes, particularly when tasks are difficult.

Numeracy and bias

The protocol analysis provided a more detailed description of the types of strategies and thus biases that influence risk reduction comparisons. The most commonly used strategy was the direct comparison of the two risk reductions (strategy S1). This indicates that the most common bias – which is particularly pronounced in low numerate individuals – is a type of *reference class neglect.* Participants tend to interpret all risk reductions as if concerning the same population. Consistent with Galesic, Garcia-Retamero, and Gigerenzer (2009), this suggests that most people are better able to understand absolute risk reduction (which always concerns the whole population) as compared to relative risk reductions (which are normalized to the population who would otherwise suffer the effect from the risk). This further suggests that people may commonly overestimate the helpfulness of a treatment when expressed as a relative risk reduction (Covey, 2007).

Individuals high in numeracy were less likely to exhibit reference class neglect: Numerate individuals performed as well in comparing two absolute as two relative risk reductions. It must be stressed that our sample was generally highly numerate, in comparison with the scores observed in the general population (cf. Galesic & Garcia-Retamero, 2008). This suggests that only a very small portion of the population would normally be likely to understand the meaning of relative risk reductions.

Results indicate that, in order to limit reference class neglect, risk reduction communications should be presented in an absolute reduction format. Relative risk reduction is an artificial construct, normalized to the base-line risk. When relying on relative risk reduction for risk communication one must be sure that the interlocutor is highly numerate and base-line risks should also be provided. Given that relative risk reductions do not provide additional information and otherwise may obscure benefits, their use should be limited.

Our study also revealed an unanticipated obstacle in the interpretation of risk communication data. Although we provided the base-line risk, made clear the reference class concerned by the risks reductions, and consistently used the same wording (frequency or single-event probability), protocol analysis revealed an unexpected interpretation of some texts and its associated normative strategy (namely $S5_b$). Generally, process tracing techniques, such as protocol analysis, seem to be valuable yet underutilized when examining biases and the rationale of various judgments.

Numeracy and complexity of strategy

Highly numerate participants consistently used more complex strategies (even in the task T_1, where the normative strategy, S1, is simple). They also used more numerical cues and used a cue (the reference class) whose relevance to the problem was not immediately obvious. This suggests that numeracy may generally predict a beneficial metacognitive

style that relies on more careful and elaborative cognitive processing (Cokely & Kelley, 2009). This also suggests that the superior decision performance in task T_2 may be partly explained by elaborative heuristic search strategies.

Frequency format and mixing of reference classes

This experiment indicates that effective, transparent risk communication should not mix absolute with relative risk reductions. Such a presentation is confusing, even for highly numerate individuals, and thus very few people manage to use a normative strategy. It must be emphasized that this cannot be attributed to a simple lack of understanding of relative risk reductions, given that many highly numerate participants performed well when comparing two relative risk reductions. The critical problem stems from the mixing of two different reference classes. One requirement of transparent risk communication is therefore the consistent use of the same reference class. Our results also suggest that the frequency format may benefit decision making as they seem to attract more attention to the difference in reference classes (although some caution is merited when generalizing these results, as only a few people gave normative answers to tasks T_3 or T_4). Generally, more empirical investigation is needed to explain when frequencies differ from single-event probability, and to test specific theoretical predictions.

The current study concerned very specific frequency and single-event probability formats, which were strongly isomorphic. Future research may also need to investigate other frequency formats, for example using integers (like "4 out of 100 persons"; see Barton, Mousavi & Stevens, 2007, for a statistical taxonomy differentiating frequency formats). One could also investigate numbers needed to treat, which is the inverse of the absolute risk reduction, rounded to the closest integer. In this case an absolute risk reduction of 4% corresponds to a number needed to treat of 25. This means that 25 people should be treated so that one would be saved. Such numbers give a good sense of which treatment is the most useful – the treatment whose number needed to treat is the lowest is the most helpful for the whole population. Moreover, it enables an estimate of the balance between benefits on one hand and side-effects (or costs) on the other hand. Is it reasonable to have 25 persons suffering side-effects so that one would be saved? Further studies should investigate which strategies people use when they need to compare numbers needed to treat from different treatments[5]. We suspect that in such a case, people would tend to perform better as compared to a task requiring the comparison of two relative risk reductions. Critically, future research and theory on the influence of context must not neglect the dynamic influences of variables such as strategies, skills, and task complexity.

---

[5] In a case, contrary to Sheridan, Pignone & Lewis (2003), where base-line risks could vary.

References

Barton, A., Mousavi, S., & Stevens, J. R. (2007). A statistical taxonomy and another "chance" for natural frequencies. *Behavioral and Brain Sciences, 30*, 255-256.

Cokely, E. T., & Kelley, C. M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgment and Decision Making, 4,* 20-33.

Covey, J. (2007). A meta-analysis of the effects of presenting treatment benefits in different formats. *Medical Decision Making, 27*, 638–654.

Edwards, A., Elwyn, G. J., Covey, J., Mathews, E., & Pill, R. (2001). Presenting risk information—A review of the effects of "framing" and other manipulations on patient outcomes. *Journal of Health Communication, 6*, 61–82.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (Revised Ed.). Cambridge, MA: MIT Press.

Galesic, M., & Garcia-Retamero, R. (2008) Statistical numeracy for health: A cross-cultural comparison. Manuscript submitted for publication.

Galesic, M., Garcia-Retamero, R., & Gigerenzer, G. (2009). Using icon arrays to communicate medical risks to low-numeracy people. *Health Psychology, 28*, 210-216.

Gigerenzer, G. (2003). Why does framing influence judgment? *Journal of General Internal Medicine, 18*, 960–961.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102*, 684–704.

Grice, H. P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.

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

Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H.G. (1997). The role of numeracy in understanding the benefit of screening mammography. *Annals of Internal Medicine, 127*, 966-972.

Sheridan, S., Pignone, M. P., & Lewis, C. L. (2003). A randomized comparison of patients' understanding of number needed to treat and other common risk reduction formats. *Journal of General Internal Medicine, 18*, 884–892.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*, 453-458.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90*, 293–315.