Journal of Experimental Psychology: Applied

How Do People Judge Risks: Availability Heuristic, Affect Heuristic, or Both?

Thorsten Pachur, Ralph Hertwig, and Florian Steinmann

Online First Publication, May 7, 2012. doi: 10.1037/a0028279

Journal of Experimental Psychology: Applied

1076-898X/12/$12.00 DOI: 10.1037/a0028279

How Do People Judge Risks: Availability Heuristic, Affect Heuristic, or Both?

Thorsten Pachur
University of Basel

Ralph Hertwig
University of Basel and Max Planck Institute for Human Development, Berlin, Germany

Florian Steinmann
University of Basel

How does the public reckon which risks to be concerned about? The availability heuristic and the affect heuristic are key accounts of how laypeople judge risks. Yet, these two accounts have never been systematically tested against each other, nor have their predictive powers been examined across different measures of the public's risk perception. In two studies, we gauged risk perception in student samples by employing three measures (frequency, value of a statistical life, and perceived risk) and by using a homogeneous (cancer) and a classic set of heterogeneous causes of death. Based on these judgments of risk, we tested precise models of the availability heuristic and the affect heuristic and different definitions of availability and affect. Overall, availability-by-recall, a heuristic that exploits people's direct experience of occurrences of risks in their social network, conformed to people's responses best. We also found direct experience to carry a high degree of ecological validity (and one that clearly surpasses that of affective information). However, the relative impact of affective information (as compared to availability) proved more pronounced in value-of-a-statistical-life and perceived-risk judgments than in risk-frequency judgments. Encounters with risks in the media, in contrast, played a negligible role in people's judgments. Going beyond the assumption of exclusive reliance on either availability or affect, we also found evidence for mechanisms that combine both, either sequentially or in a composite fashion. We conclude with a discussion of policy implications of our results, including how to foster people's risk calibration and the success of education campaigns.

Keywords: risk perception, dread, availability heuristic, affect heuristic, value of a statistical life

Health-care expenditures in developed countries are increasing at a staggering rate. In the United States, for instance, the total health-care spending in 2009 was estimated to be $2.5 trillion, accounting for 18% of the gross domestic product (see Sullivan et al., 2011). Within health-care costs, diseases contribute in different degrees to this rise. Cancer care, for instance, grew from $27 billion in 1990 to $90 billion in 2008, and is estimated to rise to $157 billion (in today's dollars) by 2020 (Sullivan et al., p. 934). There are many drivers of this explosion in the costs of cancer care, ranging from increasing cancer incidence, increased access to care, to innovation in expensive treatment procedures. A *Lancet Oncology* commission concluded that without a real effort to address the cost explosion, the "economic burden of health care in general, and high-quality cancer care in particular, will become unaffordable" (Sullivan et al., p. 951).

How to respond to this looming crisis in medical-care delivery? In a world of finite resources and competing societal priorities, we need to consider whether resources are efficiently and appropriately allocated. This process of determining cost effectiveness, the *Lancet* commission concluded, requires "socially derived assignment of value to clinical outcomes," and therefore the "input from important stakeholders such as patient groups or the general public" (Sullivan et al., 2011, p. 944). Health care is just one of myriad domains in which finite resources thrust the goals of cost effectiveness and the consideration of societal preferences and values onto policy makers. Moreover, these goals not infrequently clash. According to Sunstein (2002), too often "governments are making stabs in the dark," devoting "resources to little problems rather than big ones" and reacting to "short-term public outcries" (p. viii). When public fears rise, political pressure rises on governments to act, even when the feared disaster is very unlikely to happen or proved to be far less of a threat than had been feared (e.g., in November 2011, German health authorities announced the pending destruction of 16 million unused, expiring swine flu vaccines worth €250 million and purchased amid fears of a pandemic two years ago; Länder vernichten Millionen Dosen H1N1-Impfstoff, 2011). Similarly, governments can prove unwilling to spend frac-

Thorsten Pachur and Florian Steinmann, Department of Psychology, University of Basel, Basel, Switzerland; Ralph Hertwig, Department of Psychology, University of Basel and Max Planck Institute for Human Development, Berlin, Germany.

We thank Laura Wiles for editing the manuscript. The project was supported by the Stiftung Suzanne und Hans Biäsch zur Förderung der Angewandten Psychologie and by Swiss National Science Foundation grant 100014 125047/2 awarded to Thorsten Pachur.

Correspondence concerning this article should be addressed to Thorsten Pachur, Department of Psychology, University of Basel, Missionsstrasse 60/62, CH-4055 Basel, Switzerland. E-mail: Thorsten.Pachur@unibas.ch

tions of their limited resources on key preventive infrastructure—for instance, on levees in the years before New Orleans' destruction by Hurricane Katrina (The shaming of America, 2005)—when the public fails to be worried about a specific serious risk.

In light of the key role that the public's voice plays in democratic societies' allocation of finite resources and the delivery of legitimate solutions to diverse risks, it is important to understand the psychological mechanisms underlying risk perception. The public's risk perception affects, for instance, "the priorities and legislative agendas of regulatory bodies, such as the United States Environmental Protection Agency" and represents a key element in the "risk-assessment battlefield" more generally (Slovic, 2000, p. 390). The pioneering work of Paul Slovic and colleagues (for an overview, see Slovic, 2000) has led to a profound understanding of the dimensions of the public's risk perception. Less is known, however, about the precise cognitive and affective mechanisms involved. In this article, we aim to make progress by pitting against each other two key heuristics that people have been proposed to use in the process of judging risks.

Two Heuristics for Judging Risk: Availability and Affect

The inherently subjective construct "risk" means different things to different people. Experts' judgments correlate highly with technical estimates of annual fatalities (Slovic, 2000, p. xxiii). Laypeople's judgments about risks sometimes substantially deviate from those of experts, and they do not necessarily unfold in the form of a deliberative and analytic process; they can "be highly dependent upon intuitive and experiential thinking, guided by emotional and affective processes" (Slovic, p. xxxi). How to model these processes? Two heuristics have been proposed and studied as possible mechanisms underlying nonexperts' risk judgments: the *availability heuristic* and the *affect heuristic*. Next, we describe both of them in more detail.

Availability Heuristic

According to Tversky and Kahneman's (1973) availability heuristic,

> People assess the frequency of a class or the probability of an event by the ease with which instances or occurrences can be brought to mind. For example, one may assess the risk of heart attack among middle-aged people by recalling such occurrences among one's acquaintances. (Tversky & Kahneman, 1974, p. 1127)

The heuristic thus assumes that people infer the distal criterion (i.e., event frequency) by exploiting a proximal cue—namely, the mental availability of relevant instances. In the wake of the heuristic's conception, numerous studies have investigated the role of availability in judgment and decision making (e.g., Jacoby, Kelley, Brown, & Jasechko, 1989; Oppenheimer, 2004; Schwarz et al., 1991; Taylor & Thompson, 1982; Watkins & LeCompte, 1991; for an overview, see Betsch & Pohl, 2002, and Kahneman, 2011, Ch. 12). But the notion of availability and its potential to cause systematic judgmental biases also quickly piqued the interest of scholars of risk. Lichtenstein, Slovic, Fischhoff, Layman, and Combs (1978; see also Folkes, 1988) carried out the most influential examination of the availability heuristic in the context of risk judgments. They asked participants, for instance, to consider pairs of causes of death, say, stroke and asthma, and indicate which one causes more people to die. Across several studies, the authors found people's judgments of the frequency of various causes of death to be linked to availability. Specifically, they observed that two proxies of availability—direct suffering (e.g., at least one close friend or relative has died from the event) and indirect suffering (i.e., how often had they heard about the event via the media as a cause of death)—"did most of the job of predicting the subjects'" (p. 571) risk frequency judgments. Lichtenstein et al. also linked people's use of availability to several systematic biases in their respondents' judgments.

Building on Lichtenstein et al.'s (1978) seminal work, Hertwig, Pachur, and Kurzenhäuser (2005) operationalized different cognitive mechanisms based on the availability account and tested their specific predictions (see also Pachur, Hertwig, & Rieskamp, in press). They found that among different instantiations of availability (e.g., fluency, number of instances recalled), availability-by-recall, which relies on knowledge about relevant instances in one's social network, best predicted people's judgments of which of two risks claims more lives. Henceforth, we will adopt Hertwig et al.'s (2005) operationalization of *availability-by-recall*. It works as follows:

> *Availability-by-recall:* Recall instances of Risk A and Risk B, respectively, from your social network (encompassing family, friends, and acquaintances), and infer that risk to be more prevalent in the population for which more such instances can be recalled.

To illustrate, a person thus judges whether more people die from leukemia or from suicide by recalling actual fatalities within his or her proximate social network (see also Lichtenstein et al., 1978, Experiment 4; Fischhoff, Gonzalez, Lerner, & Small, 2005). The cause of death for which more instances can be recalled will be judged to claim more lives in the population (for the processes underlying people's search in social memory, see, Hills & Pachur, 2012).

Other key players in bringing specific risks to mind are the media, which are often skewed toward novelty, rarity, and poignancy (e.g., Eisenman, 1993; Meyer, 1990; Park & Grow, 2008). For example, a moviegoer who has just watched *Jaws* (Zanuck, Brown, & Spielberg, 1975) would likely have little trouble retrieving the (albeit fictitious) victims of a shark attack. Consistent with the media's distorting impact on our perception of the frequency of risk events, Lichtenstein et al. (1978) observed that people's median estimated frequency was higher for risks for which media coverage was more extensive. In light of the potentially important warping role that media coverage can have on the public's risk perception, we will also implement (in Study 2) a variant of availability-by-recall, *availability-by-recall*$_{TEX}$ (with TEX for Total EXperience). This mechanism works as availability-by-recall, but retrieves instances from both a person's social network *and* the media (e.g., news, novels, and movies), thus total experience.

Affect Heuristic

Images of risk often come with strong emotional reactions. Consider, for instance, the burning and collapsing World Trade

Center on September 11, 2001. Even 10 years after the event, images of it can evoke powerful and painful emotions. Indeed, risks are inextricably linked with emotions (such as dread; Slovic, 1987). Focusing on this link, Slovic and colleagues (e.g., Finucane, Alhakami, Slovic, & Johnson, 2000) proposed that "in the process of making a judgment or decision, people consult or refer to an 'affect pool'" (Slovic, Finucane, Peters, & MacGregor, 2002, p. 400). According to Slovic et al.'s affect heuristic, people may use their affective response to a risk (e.g., "How do I feel about genetically modified food, nuclear energy, breast cancer, or guns?") to infer how large they consider the risk to be (e.g., "What is the annual death toll of breast cancer?"). Consistent with the notion of the affect heuristic, there is converging evidence that emotions guide and impact risk judgments (see, e.g., Johnson & Tversky, 1983, 1984; Loewenstein, Weber, Hsee, & Welch, 2001; Pachur & Galesic, in press; Schwarz & Clore, 1983; Sjöberg, 2007; Slovic & Peters, 2006), and that different emotions can vary in their influence on risk judgments (e.g., Lerner, Gonzalez, Small, & Fischhoff, 2003; Lerner & Keltner, 2000). Here, we define the affect heuristic as follows, building on Slovic's (1987) analysis of the role of dread in risk perception:

> *Affect heuristic:* "Gauge your feeling of dread that Risk A and Risk B, respectively, evoke and infer that risk to be more prevalent in the population for which the level of dread is higher."

How can we measure the amount of dread a risk conjures up in a person's mind, thus permitting us to derive a predicted judgment for the affect heuristic? First, we linked the affect heuristic to Slovic and colleagues' psychometric framework of risk perception (e.g., Fischhoff, Slovic, Lichtenstein, Read, & Combs, 1978), in which the amount of dread evoked by a risk is captured by people's evaluations of risks on characteristics such as "voluntariness" or "certain to be fatal." Specifically, we tapped into the "affect pool" that the affect heuristic is thought to access by measuring a risk's dread score based on the characteristics that Slovic and colleagues identified as representing the "dread factor" underlying people's risk perception (see below). Furthermore, in Study 2 we measured dread by posing a single question: "Indicate the extent to which considering the cause of death triggers the feeling of dread." For both instantiations of the affect heuristic, a person infers which of two risks claims more lives by selecting the one that (based on his or her individual ratings) reaches the higher dread score or value on the dread item, respectively.[1] Finally, in Study 2 we used a third and more general approach by asking people to indicate how strongly each risk evokes each of seven basic emotions (for a similar approach, see Sjöberg, 2007): anger, disgust, fear, happiness, sadness, surprise, and contempt (e.g., Ekman & Friesen, 1986). Because this instantiation of the affect heuristic performed poorly in predicting people's responses, we do not report the results here. Note, however, that a factor analysis of the ratings on the seven emotions showed that the concept of dread proves to be closely linked to fear and sadness.

Although the availability and the affect heuristics are often listed side by side as accounts of how people assess risks (e.g., Breckenridge & Zimbardo, 2006; Keller, Siegrist, & Gutscher, 2006), we do not know of a single attempt to test these two accounts against each other. It is important to note that the heuristics may not have equal amounts of influence over people's judgments of risk. For instance, Slovic et al. (2002) hypothesized that "using an overall, readily available affective impression can be easier—more efficient—than . . . retrieving from memory many relevant examples" (p. 400). Moreover, Finucane et al. (2000) pointed out that the "challenge is to begin hypothesizing about and testing models of judgment that elucidate the roles of both cognition and affect" (p. 14). Even a decade later, however, this challenge has remained mostly unaddressed. Our goal in this article is to test—to the best of our knowledge for the first time—the relative extent to which the availability heuristic and the affect heuristic can predict people's risk judgments (for a discussion of the importance of such model comparisons, see Pachur, 2011). Moreover, by examining several measures of how people assess risks (e.g., frequency judgments, value of a statistical life)—henceforth measures of risk—we aim to carve out possible boundary conditions for the use of each heuristic in the context of risk judgments. We next turn to a description of the different measures of risk employed in our studies.

Three Measures of Risk

We employed three different measures of risk. First, we asked people to judge risks—various causes of death—in terms of their objective mortality rates (i.e., annual frequency of occurrence). *Perceived frequency* of a risk is an often-studied component of risk perception because it can be precisely explained to respondents, and their responses can be evaluated against an objective standard, thus identifying possible systematic distortions (cf. Brown & Siegler, 1993; Lichtenstein et al., 1978; Pachur & Hertwig, 2006). But perceived frequency is, of course, not all there is to people's assessments of risk. According to Slovic, Monahan, and MacGregor (2000), the public—unlike experts—sometimes draws distinctions that do not enter tabulated mortality statistics, such as that between "better deaths" and "worse deaths" (e.g., in terms of prolonged pain) or death due to an unlucky accident (e.g., lightning) or a voluntary activity (e.g., smoking). To access such distinctions and to find out how they link up with the use of the availability and affect heuristics, we employed two further risk measures. Specifically, we asked people to indicate for each risk the *value of a statistical life* (VSL), which refers to the cost of reducing the number of deaths in a specific class of risk by one. The VSL measure is used in a wide range of fields, including economics, health care, worker safety, and environmental-impact assessment. To obtain VSL judgments, we asked respondents to imagine that one single person dies from each risk each year and to indicate for each risk the amount of public money that should be invested to save the person from dying from the risk (cf. Tengs et al., 1995). It is worth pointing out that this question, unlike that of probing mortality rates, conjures the image of an individual person at risk.

Mortality statistics imply that there is an objective risk that can be measured. Slovic (2000), however, has challenged the idea that risk is objective, and emphasized its inherently subjective nature. To provide our respondents with the opportunity to express their

[1] We are thankful to Paul Slovic and Ellen Peters, who suggested these operationalizations of the affect heuristic to us.

subjective intuitions about risk (and to see which heuristic may underlie such judgments), our third measure to gauge people's risk judgments—*perceived risk*—was to instruct them (in Study 2) to assess which of two causes of death represents a "higher risk of dying from it" (cf. Slovic, Fischhoff, & Lichtenstein, 1985). Although perceived frequency and perceived risk appear to be two sides of the same coin, Slovic, Fischhoff, and Lichtenstein (1979) found them to diverge. For instance, in regard to nuclear power, respondents indicated the lowest perceived frequency but the highest perceived risk, suggesting that "lay people incorporate other considerations besides annual fatalities into their concept of risk" (p. 20). Based on further analyses, the authors proposed that characteristics of a risk such as its disaster potential—a component of the dread factor underlying risk perception—have an influence on perceived risk, but play no role in perceived frequency.

We collected responses on the three measures of risk in two modes, relative and absolute. For perceived frequency (Studies 1 and 2) and perceived risk (Study 2), participants were presented with pairs of risk events and asked to pick one of them (as in Lichtenstein et al., 1978).[2] For VSL judgments, participants provided absolute assessments (in Swiss francs); for perceived frequency, we also asked them for absolute frequency estimates, separately for each risk event.

Hypotheses

Do the relative contributions of availability and affect differ across the different measures of risk, and if so, how? For instance, does it make a difference whether a judgment refers to deindividualized statistical information, as in perceived frequency, or whether it prompts people to think of an individual person, as in VSL judgments? There is evidence that, relative to judgments of frequency, focusing on an individual case affected by a risk reduces reliance on availability (Jones, Jones, & Frisch, 1995), and instead increases the impact of the affective component evoked by a risk (Slovic et al., 2000; see also Small, Loewenstein & Slovic, 2007). A close link between case-specific information and reliance on affect is also predicted by Epstein's (1994) cognitive–experiential self theory. Finally, Tsai and Thomas (2011) found evidence that people rely more on contextual information (e.g., affect) when thinking concretely (rather than abstractly) about a problem.

As perceived frequency refers to deindividualized statistical information, we hypothesized that, when judging the frequency of a risk, people would primarily rely on availability-by-recall. VSL judgments, by contrast, induce focus on a concrete, individual case. We therefore hypothesized that the role of the affect heuristic would here be more pronounced, whereas availability-by-recall would prove less predictive of VSL judgment than of frequency judgments. A corollary of these hypotheses is that measures of availability should correlate more strongly with people's absolute frequency estimates than measures of affect, but that measures of affect should be predictive of people's VSL judgments. Concerning our third risk measure, we hypothesized that judging possible causes of death in terms of the risk of dying from them would lead to a stronger reliance on affect (relative to availability) than in judgments of risk frequencies. The reason is that this type of risk judgment appears to trigger the dread aspect of a risk (Slovic et al., 1979), thus fostering an affective evaluation.

Next we report two studies in which we presented participants with various types of risk events and asked them to judge the risks in terms of their (a) frequency of occurrence (Studies 1 and 2), (b) value of a statistical life (Studies 1 and 2), and (c) perceived risk (Study 2). Moreover, we probed each risk's availability and amount of dread evoked, allowing us to directly compare availability-by-recall and the affect heuristic. It is important to note that in constructing the test bed for examining the heuristics, we aimed to implement a representative design (cf. Brunswik, 1955; Dhami, Hertwig, & Hoffrage, 2004), encompassing a broad set of risks, instead of focusing on a small and selective one. To that end, we employed as risks all 24 types of cancer listed in official health statistics (Study 1) and a classic assorted set of risks (Study 2; cf. Lichtenstein et al., 1978). By using the full set of cancers and an assorted set of risks, we were also able to gauge how accurately heuristics can discriminate across a wide range of real-world risks.

Study 1: Availability and Affect in the World of Cancer

Method

Participants, material, and design. Thirty-three students (30 female, $M = 22.0$ years) from the University of Basel were presented with 24 types of cancer (Cancer set; Table 1; Hertwig et al., 2005). We collected, using a within-subjects design, judgments on two measures of risk (perceived frequency and VSL), and studied one measure (perceived frequency) using two response modes (pairwise choice vs. estimation). In a *choice task*, participants were presented with all 276 pairs of the cancers and judged which of two claims more lives per year in Switzerland (i.e., perceived relative frequency). Beyond course credit, participants received and lost 0.04 Swiss francs for every correct and incorrect judgment, respectively. Further, in an *estimation task*, participants estimated for each cancer its annual death toll in Switzerland (i.e., perceived absolute frequency). Finally, in a *valuation task* (asking for VSL judgments), participants received the following instruction:

> For each of the following risks, imagine that one single person dies from the risk each year. Further, for each risk there are measures that can prevent this fatality. However, implementing them is not without costs. How much money (in Swiss francs) should the government spend (at maximum) to avoid the one fatality caused by the respective risk?

We also collected information to measure affect and availability. In a risk questionnaire, people were asked to rate the 24 types of cancer on the 12 risk characteristics that Slovic, Fischhoff, and Lichtenstein (1980) found to contribute to the dread factor: voluntariness of risk, preventive control, control of severity, chronic-

[2] Note that a paired-comparison task has several strong points: Being comparative, it can help to reveal the internal mapping of events (e.g., Thurstone, 1927), and to avoid a potential weakness of absolute judgments, namely, the danger of underestimating people's ability to discriminate between magnitudes (Miller, 1956; cf. Gigerenzer & Richter, 1990). In addition, relative judgments reflect the fundamentally comparative nature of how people judge objects and events in the world (Stewart, Chater, & Brown, 2006).

Table 1
Entries in the Cancer Set; Mean Annual Mortality Rate in Switzerland, Averaged Across the Six-Year Period 1999–2004 (Bundesamt Für Statistik, 2007), Median Estimated Frequencies, Median Value of a Statistical Life (VSL), Mean Dread, and Number of Recalled Instances

	Annual mortality rate	*Mdn* frequency estimate	*Mdn* VSL	Affect (*M* dread score)	Availability (*M* number of instances in social network)
Penis cancer	10.2	40	10,000	4.65	0
Testicular cancer	17.2	109.5	25,000	4.55	0.03
Bone cancer	37.5	80	30,000	4.88	0.12
Thyroid cancer	69.3	80	30,000	4.44	0.03
Larynx cancer	94.2	50	20,000	3.92	0.24
Cancer of the connective tissue	94.3	45	17,500	4.4	0
Cancer of the gall bladder	196.5	30	11,010	4.31	0.09
Malignant melanoma (skin cancer)	242	50	17,102	3.59	0.24
Cervical cancer	295.8	95	20,000	4.39	0.15
Renal cancer	339.2	50	20,000	4.3	0.03
Cancer of the mouth and throat	351	60	12,609	3.91	0
Esophageal cancer	384.5	50	15,000	4.31	0.09
Rectal cancer	437.2	80	10,000	4.14	0
Bladder cancer	450.5	30	10,616	4.23	0.03
Ovarian cancer	453.2	100	20,000	4.43	0.03
Cancer of the nervous system	455	90	30,000	5.02	0.32
Hepatic cancer	513	110	10,000	4.04	0.18
Stomach cancer	572.2	115	20,000	4.22	0.18
Pancreatic cancer	897.8	60	17,501	4.43	0.26
Colon cancer	1,172.2	120	10,000	4.23	0.15
Prostate cancer	1,312.3	87	23,115	4.53	0.32
Leukemia and lymphoma	1,331.7	100	35,000	4.91	0.38
Breast cancer	1,347.3	200	40,000	4.24	0.79
Lung cancer	2,756	300	22,500	4	0.5

catastrophic, common-dread, certain to be fatal, equity of risks and benefits, threat to future generations, personal exposure, potential of global catastrophe, changes in risk, and ease of reduction. Participants rated risks on these characteristics using a 1–7 scale. For example (cf. Slovic et al., 1985): "Is this a risk that kills people one at a time (chronic risk) or a risk that kills large numbers of people at once (catastrophic risk)?"

Finally, in a *recall task*, participants reported for each type of cancer the number of deaths that they could recall from their social network.

Procedure. The choice task was presented on a computer; all other tasks were administered as paper-and-pencil questionnaires. Participants completed all tasks in the following order: choice task, risk questionnaire, estimation task, recall task, and valuation task.[3] The order in which the risks appeared in the individual tasks was randomized for each participant (as were the risks within each pair in the choice task). Between tasks, participants were encouraged to take a break whenever necessary, to avoid fatigue. Completing all tasks took between 75 and 90 min.

Results

In the choice task, participants made, on average, 62.2% correct inferences. For the estimation task, Table 1 shows the median estimated frequency for each risk. The correlation between the estimates and the actual frequencies (both log-transformed to reduce skew) was substantial, although not very high, $r = .46$ ($p = .02$; mean individual $r = .26$). Table 1 shows that in the valuation task, people did not simply assign the same monetary values to all types of cancer. To illustrate, for leukemia, the median VSL was 35,000 Swiss francs, relative to 11,010 francs for cancer of the gall bladder. The dread score for each risk was calculated as the mean rating on the 12 characteristics assessed in the risk questionnaire, coded such that a higher value indicates higher dread. (The risks' factor scores, representing a weighted mean, yielded qualitatively very similar results.) Across risks, the mean dread score and the (log-transformed) mean number of recalled instances (see Table 1) were uncorrelated, $r = .04$ ($p = .85$; the mean individual $r = .03$), a good test bed for pitting availability-by-recall and the affect heuristic against each other.[4]

How well do availability-by-recall and the affect heuristic predict relative judgments of risk? In order to derive individual- and item-specific predictions of availability-by-recall and the affect heuristic in the choice task, for each participant and each pair of risks, we determined for which of the two risks the participant recalled more instances, and which risk the person rated as triggering more dread (based on the dread score), respectively. By employing the full set of cancers (rather than focusing on

[3] One may suspect order effects. Judging from the very similar results for the choice and estimation tasks (administered as first and third task, respectively), however, such effects appear negligible. Moreover, in Study 2, where the task order was counterbalanced, we found very similar results to those in Study 1.

[4] Given the high proportion of female participants, one may suspect that the results could be distorted by breast cancer, for which participants recalled the largest number of instances (see Table 1). Additional analyses showed, however, that the results and the conclusions from the statistical tests are not affected when breast cancer was excluded from the analyses.

selected pairs), we were able to evaluate how often availability-by-recall and the affect heuristic could discriminate between the risks. The heuristics failed to discriminate when a person recalled the same number of instances or no instances for both risks or when both risks triggered in the eye of the person the same amount of dread. Across participants, availability-by-recall discriminated in, on average, 69.4 (25.2%) pair comparisons and the affect heuristic in 176.0 (63.8%) pair comparisons. In other words, people's affective responses to risks enabled markedly more discriminations than their knowledge of risk occurrences in their social network.

When comparing the descriptive accuracy of availability-by-recall and the affect heuristic, it is important to take into account these differences in the heuristics' power to discriminate. For instance, one heuristic may reach a lower overall performance than the other, merely because it makes predictions for a larger set of items, including "difficult" ones, whereas the other renders predictions for relatively "easy" items only. Therefore, in the following analysis we focused on the subset of pair comparisons in which *both* heuristics discriminated. This was the case, on average, for 63 items per participant (23.1% of all comparisons) in the choice task.

Figure 1 shows the mean (across participants) percentage of correct predictions for availability-by-recall and the affect heuristic, separately for the relative frequency judgments (i.e., the choice task) and the VSL judgments. In the relative frequency judgments, availability-by-recall outperformed the affect heuristic with a lead of 19.6 percentage points, $M = 74.0$ ($SD = 12.9$) versus 54.4% ($SD = 16.1$), $t(28) = 5.07$, $p = .001$, $d = 1.35$. Furthermore, only 13.8% of participants were better predicted by the affect heuristic than by availability-by-recall. What about the VSL judgments? To be able to conduct a pair-comparison analysis for these, we used, separately for each person, responses in the valuation task (where they provided absolute judgments) to simulate choices in all 276 pairs of cancers. Had the person, for example, indicated a higher VSL for leukemia than for cancer of the gall bladder, we assumed that he or she would have *chosen* to spend more on preventing the former (items where both cancers were valued equally were not considered in this analysis; therefore, for the VSL judgments the number of items on which both heuristics made predictions equaled, on average, 34 items per participant, or 12.3% of all comparisons). We then tested these simulated VSL choices against the predictions of availability-by-recall and the affect heuristic. Contrary to the choices regarding perceived frequency, Figure 1 shows that availability-by-recall and the affect heuristic here performed indistinguishably, $M = 60.3\%$ ($SD = 26.6$) versus $M = 61.9\%$ ($SD = 23.8$), $t(20) = 0.19$, $p = .86$, $d = -0.06$. Moreover, now about half of participants (52%) were better predicted by the affect heuristic than by availability-by-recall.

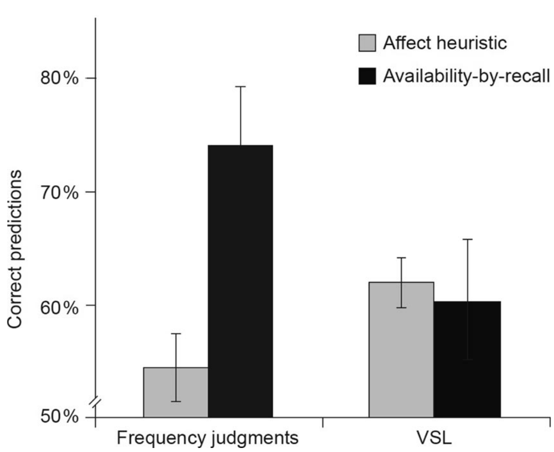

Figure 1. Percentage of correct predictions of the affect heuristic and availability-by-recall for the frequency choices and the simulated VSL (= value of a statistical life) choices in Study 1. Error bars represent ± 1 standard error of the mean.

The results in Figure 1 suggest that the heuristics' descriptive validity depends on the risk measure used. This dependency was corroborated by a repeated-measurement ANOVA (with heuristic and type of risk measure as within-subjects factors), showing a significant interaction between heuristic and type of risk measure, $F(1, 20) = 8.16$, $p = .01$, $\eta_p^2 = .29$. To wit, the impact of affective information appears substantially more pronounced in monetary evaluations of the value of an individual life than in judgments of mortality frequencies; the latter kind of judgments are well explained in terms of recall of direct experiences. As pointed out earlier, availability-by-recall does not discriminate between risks as frequently as the affect heuristic. Therefore, one could ask whether the latter may step in to help when the former does not render possible a decision. The brief answer is yes and the respective analyses can be found in Appendix A.

How well do availability and affect predict absolute judgments of risk? The descriptive analysis of the heuristics in relative judgments reported above did not consider cases where the heuristics failed to discriminate between two risks, thus excluding some risks. How robust are the conclusions with regard to people's reliance on availability and affect when considering all risks? To answer this question, we analyzed how the measures of availability and affect were related to participants' responses in the frequency estimation task and the valuation task, which involved absolute judgments of individual risks rather than pairs of risks. Both correlation and regression analyses were conducted. For comparability with Lichtenstein et al.'s (1978) analyses, we examined how aggregate frequency estimates and VSL judgments for the different risks were related to aggregate measures of availability and affect. To this end, the number of recalled instances, the estimated frequencies, VSL judgments, and actual frequencies (see Table 1) were log-transformed to reduce skew in the distribution.

Replicating results by Lichtenstein et al. (1978), the median (across participants) estimated frequency for a risk was strongly related to the mean number of recalled instances, $r = .62$, $p = .001$. (Lichtenstein et al. reported correlations between $r = .50$–.90). Available risks, that is, risks for which participants recalled many instances from their social network, were estimated to be more frequent in the population than less available risks. In contrast, the mean amount of affect triggered by a risk (i.e., dread score) was orthogonal to estimated frequency, $r = .04$, $p = .87$. Did affect hold more sway over the VSL judgments? Indeed, median VSL judgments were related to both availability, $r = .56$, $p = .005$, and affect, $r = .44$, $p = .03$, suggesting again that the role of affect differs between different risk measures. Note that, as shown in Table 1, the lower correlations for affect were not due to constrained variability of the dread score, which showed clear differentiation between the risks. Analyses based on aggregate data can be tainted by aggregation bias (e.g., Cohen, Cohen, West, & Aiken, 2003, p. 539). Therefore, in Appendix B, we also report the results on the individual level. Although, due to the larger amount of unsystematic variance in individual data, the correlations were lower than for the aggregate analyses, the same pattern emerged.

Because we did not manipulate availability experimentally (but gauged people's experiences in their social networks), the correlation between frequency estimates and availability could be spurious, caused by their shared variance with the objective mortality statistics. This, however, did not turn out to be the case. Availability and frequency estimates were correlated even when controlling for actual frequency, $r = .48$, $p = .02$.

Finally, to examine the independent effects of availability and affect on people's risk judgments, we conducted regression analyses. Specifically, absolute frequency estimates and VSL judgments for a risk were regressed on affect and availability, both on the aggregate and the individual level. With regard to the aggregate analyses (i.e., predicting median frequency estimates and VSL judgments based on the each risk's mean score on the availability and affect measures), Table 2 shows that although availability was the strongest predictor for both frequency estimates and VSL judgments (i.e., it had the highest regression coefficients), there were important differences. For the absolute frequency estimates, affect was not a significant predictor, whereas for the VSL judgments availability and affect contributed about equally. In Appendix B, we report that this pattern was also obtained when regression models were computed for each individual participant. In sum, these results suggest that whereas affect has only a weak impact in frequency estimates of the risks, it plays a clearly larger role in VSL judgments.

How ecologically valid are availability and affect? Using the mortality rates in Switzerland (averaged across the 6-year period 1999–2004; e.g., Bundesamt für Statistik, die Schweiz, 2007) as a reference, we also examined the ecological accuracy of availability-by-recall and the affect heuristic, based on the items in which they discriminate between risks. The heuristics' ecological validity, that is, the percentage of accurate inferences given that the heuristics discriminated, is shown in Table 3. Availability-by-recall achieved an accuracy of 71.4% ($SD = 18.0$), whereas the affect heuristic did not perform better than chance (48.7%, $SD = 17.3$).

Table 3
Percentage of Cases in Which the Different Operationalizations of the Affect and Availability Heuristics Pointed to the Risk With the Higher Frequency (i.e., Mortality Rate)

	Ecological accuracy for judging the risks' actual frequencies	
	Study 1	Study 2
Affect heuristic (dread score)	47.6	70.2*
Affect heuristic (dread item)	—	76.6*
Availability-by-recall	71.9*	81.5*
Availability-by-recall$_{TEX}$	—	69.0*

* Accuracy is significantly better than chance ($p < .05$).

Summary

In Study 1, we found that when responding to an exhaustive set of cancers, people appear to recruit availability-by-recall and the affect heuristic to different degrees, depending on the measure of risk. When judging relative mortality frequencies (which of two cancers claims more lives?) and absolute frequencies (what is the annual death toll of a cancer?), people appear to primarily rely on the number of recalled instances from their social network. When judging how much money should be spent to avoid a single death due to a specific cancer, by contrast, a substantial portion of people appear to consult their affective response (dread) evoked by the risk.

Study 2: Availability and Affect in a Classic Set of Risks

In order to test the generalizability of our findings in Study 1, we next tested instantiations of the availability heuristic and the affect heuristic in the context of a classic set of risks: an assorted set of causes of deaths investigated by Lichtenstein et al. (1978). Unlike the cancers used in Study 1, this set represents a heterogeneous collection of causes of death, including risks such as accidents, natural hazards, diseases, and homicide. Furthermore, we asked one group of participants to judge the causes of death in terms of their perceived risk (i.e., the subjective risk of dying from them). Finally, in order to test availability-by-recall$_{TEX}$ we extended our measures of availability to instances from the media, and we also made use of an alternative, simpler measure of affect.

Table 2
Results of the Regression Analyses on the Aggregate Level

Predictor	Study 1: Frequency estimates (log)	Study 1: VSL (log)	Study 2: Frequency estimates (log)	Study 2: VSL (log)
Dread score	.02	**.41**	.02	−.20
Dread item	—	—	**.43**	**.82**
Direct experience (log)	**.48**	**.48**	**.66**	.24
Indirect experience (log)	—	—	**−.18**	−.15
R^2_{adj}	.23 ($p = .07$)	.42 ($p = .003$)	.84 ($p = .001$)	.71 ($p = .001$)

Note. In the analyses, the log-transformed median frequency estimates and value-of-a-statistical-life (VSL) judgments for the different risks were regressed on the different measures of affect (mean dread score and dread item) and availability (mean number of instances in social network = direct experience, number of instances in media = indirect experience). Significant standardized regression coefficients ($p < .05$) are in bold.

Method

Participants, material, and design. Eighty-five students (62 female; mean age = 24.9 years) from the University of Basel were presented with 41 causes of death (Assorted set*;* Table 4; Hertwig et al., 2005; Lichtenstein et al., 1978). In the choice task, one group of participants (n = 45; the *frequency-choice group*) was instructed to judge which of two causes of death has a higher annual mortality rate in Switzerland (the same task as used in Study 1); another group (n = 40; the *risk-choice group*) was instructed to judge which of two causes of death represents "a higher risk of dying from it" in Switzerland. Both groups saw the same 106 pairs used by Lichtenstein et al. (1978). Subsequently, all participants worked on the same tasks used in Study 1 (i.e., estimation task, valuation task, risk questionnaire, and recall task), with two modifications. In the recall task, in addition to reporting the number of instances in their social environment (*direct experience*) participants now also reported, separately for each risk, the number of specific instances that they could recall from sources such as the news, movies, and novels (*indirect experience*). In the risk questionnaire, dread was additionally measured using a single item,

Table 4
Entries in the Assorted Set; Mean Annual Mortality Rate in Switzerland, Averaged Across the Six-Year Period 1999–2004 (Bundesamt Für Statistik, 2007), Median Estimated Frequencies, Median Value of a Statistical Life (VSL), the Mean Responses on the Measures of Dread, and the Mean Number of Recalled Instances From the Social Network (= Direct Experience) and the Media (= Indirect Experience), Respectively

				Affect		Availability	
						M number of recalled instances	
	Annual mortality rate	*Mdn* frequency estimate	*Mdn* VSL	*M* dread score	*M* dread item	Direct experience	Indirect experience
Flood	0	15	10,000	4.45	2.81	0.02	126.42
Polio	0	25	10,000	3.47	2.12	0.08	0
Measles	0	15	5,000	3.18	1.76	0.09	0
Smallpox	0	10	3,000	3.22	1.76	0.06	0
Smallpox vaccination	0	3	4,500	2.66	1.73	0.07	0
Poisoning by vitamins	0	10	2,000	2.93	1.84	0	0.01
Whooping cough	0.2	20	5,000	3.41	2.11	0.05	0
Pregnancy, childbirth, and abortion	0.2	30	10,000	3.00	2.68	0.18	0.19
Lightning	0.3	5	1,000	4.07	2.62	0.02	0.75
Fireworks	0.3	7	2,000	3.05	2.36	0.04	0.26
Syphilis	1	12	5,000	3.11	2.07	0	0
Tornado	1.3	0	500	4.55	2.16	0	13.98
Nonvenomous animal	1.7	12	2,000	3.18	2.21	0.01	0.87
Botulism	2.8	34	5,000	3.40	2.51	0.04	0.31
Motor vehicle-train collision	3.2	20	10,000	3.97	3.82	0.19	1
Firearm accident	3.2	37	10,000	3.88	2.91	0.07	2.93
Venomous bite or sting	3.5	20	5,000	3.52	3.25	0.01	0.23
Electrocution	6.8	20	2,000	3.71	2.85	0.09	0.22
Excess cold	9	10	4,000	3.48	2.69	0.04	3.28
Infectious hepatitis	11.3	50	10,000	3.52	3.09	0.24	0.04
Appendicitis	15.7	50	5,000	3.45	2.55	0.05	0
Poisoning by solid or liquid	22.2	50	7,000	3.57	2.56	0.01	0.79
Tuberculosis	23.2	20	10,000	3.68	2.54	0.07	0.04
Fire and flames	31.8	40	5,000	4.41	3.53	0.25	18.30
Emphysema	52.5	50	6,000	3.83	3.00	0.06	0.02
Drowning	61.8	26	4,000	3.67	3.41	0.36	12.88
Homicide	66.7	50	9,000	4.75	3.53	0.20	80.19
Asthma	145	50	7,000	3.37	2.05	0.15	0
Leukemia	494.3	200	50,000	4.73	4.60	0.26	0.59
Motor vehicle (car, truck, or bus) accidents	554.3	500	20,000	4.35	5.15	1.92	96.49
Accidental falls	1,096.2	200	10,000	4.06	3.78	0.49	8.48
Suicide	1,333.7	300	6,000	3.56	2.09	1.27	6.31
Breast cancer	1,347.3	300	50,000	4.53	4.71	0.86	2.23
Diabetes	1,632.5	120	15,000	3.84	3.32	0.29	0.04
All accidents	2,094	1,000	20,000	4.52	4.44	1.59	94.46
Lung cancer	2,756	500	50,000	4.49	5.14	0.35	1.27
Stroke	2,792.5	500	30,000	4.56	4.29	0.87	1.16
Cancer of the digestive system	4,366.7	500	70,000	4.73	5.02	0.61	0.32
All cancer	15,338.2	2,000	200,000	4.96	5.58	2.84	15.85
Heart disease	24,095.2	1,000	25,000	4.59	4.89	1.78	2.60
All disease	55,507.7	7,000	100,000	4.75	4.89	4.78	50.52

asking participants simply to indicate the amount of "dread elicited when considering the risk" (on a 1–7 scale). Participants received course credit or 30 Swiss francs as compensation.

Procedure. All tasks were presented on a computer. The choice task was always presented first; the order of the other tasks was counterbalanced. In all other respects, the procedure was identical to that in Study 1. Completing all tasks took between 85 and 110 minutes.

Results

In the choice task, participants selected the risk with the objectively higher mortality rate in, on average (across participants), 69.0% (SD = 4.7) of the cases in the frequency-choice group and in 67.2% (SD = 7.3) of the cases in the risk-choice group. Table 4 shows the individual risks' median estimated frequencies and VSL judgments in the estimation and valuation tasks, respectively. The median estimated frequencies were strongly correlated with the actual frequencies (both log-transformed), r = .92 (p = .001; mean individual r = .68).

Before testing the heuristics, let us first turn to the additional measure of availability, the number of recalled instances in the media (i.e., indirect experience). How is it related to actual mortality statistics? For each risk, we subtracted its rank based on number of recalled instances in the media from its rank according to mortality statistics (both ranked in descending order). Consistent with the analysis by Combs and Slovic (1979), the mean difference score was positive for accidents (M = 7.1), natural hazards (M = 20.5), and homicide (M = 11), suggesting overrepresentation of these dangers in the media, and negative for diseases (M = −7.4) and suicide (M = −1), suggesting underrepresentation. Table 5 shows that, overall, the mean number of instances recalled from the media (log-transformed to reduce skew in the distribution) was related to the mortality statistics, although not very strongly (cf. Lichtenstein et al., 1978).

Table 5 shows that, unlike in Study 1 (cancer set), on the aggregate level (i.e., based on each risk's mean score on the measures of availability and affect), availability and affect were correlated across the events in this assorted set of risks (r = .46–.73). Table B2 in Appendix B shows that this also holds on the individual level. Affect (as measured by the single dread item) and availability (as measured by direct experience) showed the highest zero-order correlations with the actual mortality rates.

How well do availability-by-recall and the affect heuristic predict relative judgments of risk? As in Study 1, we derived, in the choice task, predictions of the different implementations of availability-by-recall and the affect heuristic for each participant and each pair of risks based on the individuals' responses on the measures of availability and affect. Across the different implementations of the heuristics, discriminability between risks in the pair comparisons varied markedly (see Table A1 in Appendix A). The affect heuristic based on the dread score and the dread item discriminated between the risks in, on average, 102.1 (96.3%) and 75.2 (71.0%) pair comparisons, respectively. Availability-by-recall (only direct experience) and availability-by-recall$_{TEX}$ (i.e., direct *and* indirect experience) discriminated in 30.7 (29.0%) and 67.4 (63.6%) pair comparisons, respectively.

In order to examine the performance of the heuristics in predicting people's choices, we focus on those pair comparisons for which all heuristics make a prediction (as done in Study 1). This was the case for, on average, 20.3 items (= 19.1% of all comparisons), 20.6 items (= 19.4%), and 17.0 items (= 16.1%), for the frequency, risk, and (simulated) VSL choices, respectively (we simulated the responses for the VSL choices using the same approach as in Study 1). Figure 2 (upper panel) shows that for judging which of two risks claims more lives, availability-by-recall predicted results much better than availability-by-recall$_{TEX}$. More important, availability-by-recall, with a performance of 81.5% (SD = 13.7) correctly predicted choices, also outperformed the best-performing instantiation of the affect heuristic (based on the single dread item; 76.7%, SD = 14.7), $t(41)$ = 2.3, p = .03, d = .34. This replicates Study 1.

Also, as in Study 1, the picture looks different for the (simulated) VSL choices. As Figure 2 (middle panel) shows, the best-performing instantiation of the affect heuristic (based on the single dread item) tied with availability-by-recall: 75.1% (SD = 16.8) versus 77.6% (SD = 14.2), $t(71)$ = −1.2, p = .25, d = −.16. Similarly, Figure 2 (lower panel) shows that for judging which of two causes of death represents a higher perceived risk, the best-performing instantiation of the affect heuristic (based on the dread score) tied with availability-by-recall, 76.6% (SD =

Table 5
Intercorrelations Between the Different Measures of Availability and Affect, Their Correlation to the Actual Frequencies of the Risks and to the Frequency Estimates and Value-of-a-Statistical-Life (VSL) Judgments in Study 2

	Availability		Affect	
	Direct experience (log)	Indirect experience (log)	Dread score	Dread item
Direct experience (log)	—	**.55**	**.60**	**.73**
Indirect experience (log)		—	**.63**	**.46**
Dread score			—	**.83**
Actual frequency (log)	**.82**	**.38**	**.71**	**.83**
Estimated frequency (log)	**.87**	**.38**	**.64**	**.82**
VSL (log)	**.71**	.26	**.62**	**.82**

Note. Significant correlations (p = .05; two-tailed) are in bold.

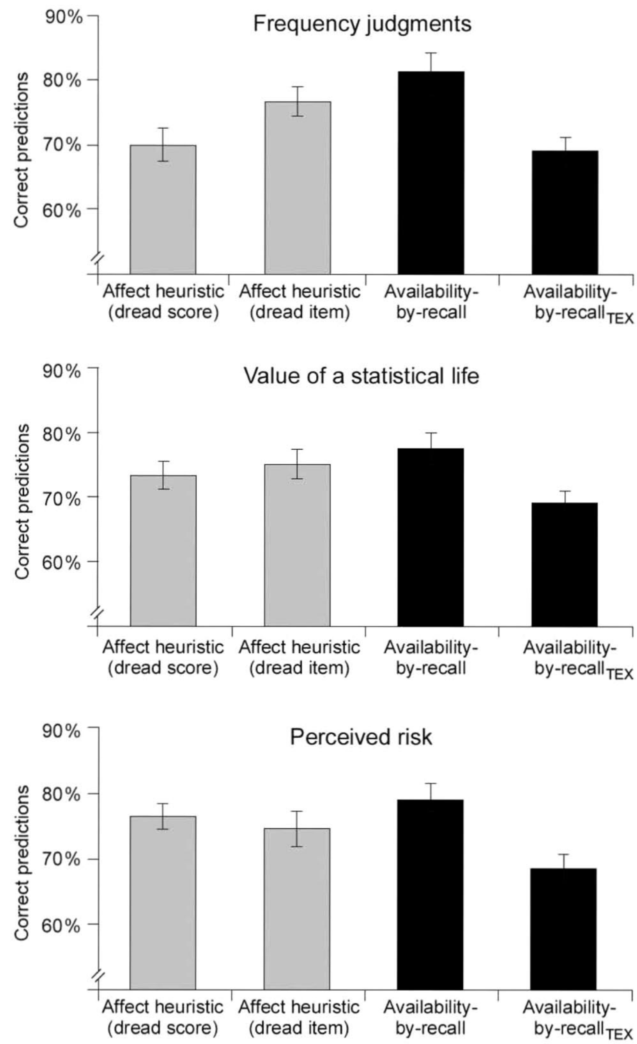

Figure 2. Percentage of correct predictions of the two instantiations of the affect heuristic, availability-by-recall, and availability-by-recall$_{TEX}$ in Study 2's choice tasks: frequency (upper panel), value of a statistical life (middle panel), and perceived risk (lower panel). Error bars represent ± 1 standard error of the mean.

12.4) versus 78.9% (SD = 14.6), $t(36) = -0.97$, p = .34, d = −.17. The dependence of the relative performance of the heuristics on the type of risk measure was corroborated by repeated-measurement ANOVAs using type of heuristic (focusing on the two best-performing heuristics: availability-by-recall and the affect heuristic based on the dread item) and type of risk measure (frequency vs. VSL and perceived risk vs. VSL in the frequency-choice and risk-choice groups, respectively) as between-subjects factors. As in Study 1, there was an interaction between the two factors for the frequency-choice group, $F(1, 39) = 4.26$, $p = .046$, $\eta_p^2 = .099$. For the risk-choice group, by contrast, there was no interaction, $F(1, 31) = 0.12$, $p = .73$, $\eta_p^2 = .004$. This indicates that availability and affect play different roles in frequency judgments versus VSL judgments, but not in VSL versus perceived-risk judgments (i.e., which of two causes of death represents a higher risk of dying from it). The same pattern across the three risk measures also emerged when analyzing the percentages of participants that were best described by either model (see Table 6).

In sum, we obtained two major results. First, when judging which of two risks claims more lives, people are more likely to rely on direct experience than on indirect experience or affective information. Second, the instantiation of the affect heuristic based on a single dread item predicts frequency choices better than the instantiation based on the more complex dread score, and ties with it for the VSL judgments and perceived-risk choices. (Appendix A reports an analysis of the two implementations of the affect heuristic and availability-by-recall$_{TEX}$ in those cases where availability-by-recall failed to discriminate.)

Could availability and affect substitute for each other? So far, we have assumed that the availability and affect heuristics are two distinct heuristics. However, this need not be so and some results suggest that (some) people may draw on both of them. How could this be modeled? One possibility is that availability and affect do not operate as distinct heuristics but represent cues and can thus function vicariously (e.g., Brunswik, 1952). Such a cue-substitution process could be modeled in terms of a cue hierarchy, in which availability and affect are processed sequentially, and the second cue substitutes the first when it fails to discriminate. Below, we test two cue-substitution processes: one where affect (measured in terms of the single dread item) substitutes direct experience if the latter does not discriminate, and one where direct experience substitutes affect if the latter does not discriminate. Both processes assume that if neither cue discriminates, people guess. A second possibility to model people's reliance on both availability and affect is to assume that judgments are based on a composite of direct experience and affect. In order to determine this composite, we z-transformed, separately for each participant, the number of instances and the responses on the dread item, summed these two z-scores for each risk, and then determined for which risk the sum was higher. If the composite yields the same value for both risks, a decision is reached by guessing.

Using the frequency choice data of Study 2, we determined which of these integrative models—the two cue-substitution processes or the composite process—performs best in predicting

Table 6
Percentage of Participants That Were Best Described by Availability-by-Recall and the Best-Performing Instantiations of the Affect Heuristic for the Frequency-Choice Task, the Simulated Value-of-a-Statistical-Life (VSL) Choices, and the Perceived-Risk Choice Task in Study 2

Choice	% of participants best described by			
	Availability-by-recall	Affect heuristic	N	p (sign test)
Frequency	57.1	23.8[a]	42	.02
VSL	45.8	34.7[a]	72	.36
Perceived risk	54.1	32.4[b]	37	.22

Note. The percentage of participants where the two best models tied was 19.1%, 19.4%, and 13.5% in the frequency, VSL, and perceived-risk choices, respectively.
[a] based on the single dread item. [b] based on the dread score.

people's judgments. As benchmark, we also calculated the performance when assuming that people exclusively rely on a one-reason heuristic—availability-by-recall or the affect heuristic (based on the single dread item), depending on which accounted for a given participant's choices best—and guess if the heuristic does not discriminate. We determined which of the possibilities predicts each person's judgments best, and then, across all people, determined the overall winner. Two mechanisms best described the largest percentage of participants: the cue-substitution process that operates on direct experience first and then affect (35.3%), and the mechanism that integrates instance knowledge and affect into a composite (32.2%). The scores for these two mechanisms were indistinguishable ($z = .42$, $p = .68$), but they surpassed the score for the cue-substitution process that considers affect first (16.7%; $z > 2.2$, $p < .03$), and that when assuming that people exclusively rely on availability-by-recall or the affect heuristic (15.9%; $z > 2.3$, $p < .02$).

How well do availability and affect predict absolute judgments of risk? In order to find out how well measures of availability and affect predict absolute judgments as rendered in the estimation task and the valuation task, the frequency estimates, the VSL judgments, and the number of instances recalled from people's networks and from the media were log-transformed to reduce skew in the distribution (cf. Study 1). Table 5 shows, using zero-order correlations, how the frequency estimates and the VSL judgments were related to the measures of availability and affect on the aggregate level. (The correlations on the individual level, which largely showed the same pattern as on the aggregate level, are reported in Appendix B.) Following Study 1 and Lichtenstein et al. (1978), the frequency estimates were correlated with direct experience (i.e., the number of recalled instances). Indirect experience (i.e., the number of instances recalled from the media) proved to be only weakly related to the estimates. The estimates were also related to the measures of affect, but less strongly than to direct experience. The VSL judgments, by contrast, were equally strongly linked to affect (as measured by the single dread item) and to direct experience. As in Study 1, the correlation between frequency estimates and direct experience also held when controlling for actual frequency, $r = .50$, $p = .001$.

Like in Study 1, we used multiple regression analyses to determine the independent contribution of the different measures of affect and availability on absolute judgments of risk. Table 2 shows the results on the aggregate level; the results for the regression analyses on the individual level are reported in Appendix B. The frequency estimates were best predicted by direct experience (as indicated by higher regression coefficients). The VSL judgments, by contrast, were best predicted by affect as measured by the single dread item. In sum, Study 2 replicates Study 1 by showing that the roles of availability and affect differ between the two risk measures, with affect playing a stronger role in the VSL judgments than in the frequency estimates.

How ecologically valid are availability and affect? We also examined the ecological accuracy of the different instantiations of the availability and the affect heuristics. For this classic, heterogeneous set of risks, we found that the ecological validity of availability and affect is much above chance across all instantiations (see Table 3).

General Discussion

What are the mental mechanisms underlying people's risk judgments? The goal of this article was to pit against each other, to the best of our knowledge for the first time, two prominent heuristics for judging risks. For that purpose, we proposed several clearly defined instantiations of the availability heuristic and the affect heuristic. Concerning the availability heuristic, we made a distinction between directly and indirectly experienced instances, thus enabling a test of which of the two drives people's judgments of risk more. Our results suggest that a person's encounters with instances of risk events that are conveyed through the mass media, relative to experience rooted in a person's proximate social network, played at best only a minor role in judgments of risk. For instantiating the affect heuristic, we proposed linking the affect heuristic with the "dread" concept in psychometric risk research (Slovic, 1987). Across Studies 1 and 2, we found that affect associated with risks can be conveniently and parsimoniously gauged by a single item asking for the amount of dread a risk evokes. This is a very useful observation for future studies on the role of affect in risk judgments.

Across two studies involving a large assortment of different risks—both common and less common ones—availability-by-recall offered a substantially better descriptive account than the affect heuristic when people judged deindividualized, statistical mortality rates. Affect, however, was at least on par with availability when people were asked to put a price tag on a single life saved from a risk, or when they were asked to indicate the perceived risk of dying from a cause of death. One possible moderator for the relative influence of availability and affect could thus be the degree to which the measure used to probe people's risk perception invokes the image of an individual (cf. Slovic et al., 2000).

In sum, these findings advance our understanding of the mental mechanisms underlying risk judgments in several ways: first, by disentangling the role of different measures of availability and affect for predicting risk judgments; second, by evaluating the relative merits of two key heuristics of risk judgments; and third, by helping to carve out the boundary conditions of the use of availability-by-recall and the affect heuristic, respectively. Although there seems to be a primacy of availability (direct experience) in risk judgments, we identified two situations in which affective information gains influence: when instance knowledge does not discriminate (Appendix A) and when the specific risk measure conjures up the image of an individual case. In what follows, we discuss potential objections to our results, the role of media on risk judgments, limitations of our studies, as well as policy and practical implications.

People's (Fortunately) Limited Experience With Risks and Cue Substitution

One objection to our test of the availability heuristic is that, due to the representative sample of risks that we used (e.g., in Study 1: all cancers, including rare ones), respondents were unable to consistently retrieve multiple instances of the respective risks. In fact, for 17% (Study 1) and 7% (Study 2) of the risks, participants could not retrieve any instances at all (Tables 1 and 4), resulting in a relatively low discrimination rate of availability-by-recall. In our view, this finding is not a weakness of our investigations but tells us something important about the risk ecology in the social world. Specifically, the low discrimination rate reveals a price that the use of the best-performing implementation of the availability heuristic in our studies exacts: Availability-by-recall cannot be recruited across the whole range of pairs of risks (but note that limited knowledge does not necessarily lead to poor judgments; Pachur, 2010). Accordingly, when judging the risks' objective frequencies, people's accuracy dropped substantially when their direct experience failed to discriminate relative to when it discriminated, both in Study 1, *M*s = 60.0% versus 68.8%, $t(29) = -3.5$, $p = .002$, $d = .91$, and in Study 2, *M*s = 60.6% versus 75.1%, $t(44) = -5.7$, $p = .001$, $d = 1.27$.

The restricted ability of availability-by-recall to discriminate speaks to a general feature of people's experience with risks and their risk perceptions. Happily, occurrences of risks tend to be relatively rare. Therefore, decision makers often lack immediate experience with relevant instances. Of course, as experimenters we could have employed a selective set of risks for which participants can typically recall many instances (e.g., breast cancer, heart attacks, and automobile accidents). Such a focus on common risks, however, would tell us little about how often the candidate heuristics can be employed across the full set of risks, and how people possibly substitute one heuristic for another.

When direct experience fails to discriminate, people can take (at least) two routes: They could extend their search space in memory to a *virtual circle*; that is, their encounters with death and diseases featured in the media. Alternatively, they could turn to their affective responses to the risks in question. Concerning the first route, we found little evidence that people rely heavily on instances drawn from such a virtual circle (see Table 2); on a related note, Hertwig et al. (2005) found little evidence that people make use of the mental fluency of risks as measured by the amount of media coverage to infer their frequency. Concerning the second route, we found in Study 2 that the assumption that people first retrieve direct experience and, if it fails to discriminate, exploit their affective responses, described more than a third of our respondents best. Another third of respondents were best described in terms of a composite measure that integrates direct experience and affect. Future investigations should shed more light on the sequential and simultaneous interplay of availability and affect. This interplay could also be investigated using participants who, relative to our young adults, are likely to have richer instance knowledge of various health risks. For instance, in judging fatal risks, to what extent do older people rely on their extensive instance knowledge or their affective responses to ever-more imminent risks?

Availability, Media, and Judgments of Risk

In their original description of the availability heuristic, Tversky and Kahneman (1973) did not distinguish between availability stemming from direct or indirect experience. In fact, it has often been explicitly assumed that judgments of risk may be equally driven by both (e.g., Lichtenstein et al., 1978). Our results, by contrast, show that it is important to distinguish between the two sources of availability. Specifically, whereas direct experience plays a key role in shaping people's risk judgments, we obtained no evidence that availability through instances learned from the media has an impact, once direct experience and affect are taken into account (for similar conclusions, see Hertwig et al., 2005). This contradicts the thesis that, because the media sometimes stress highly improbable risks, media coverage is instrumental in causing distorted judgments of risk (Combs & Slovic, 1979; Lichtenstein et al., 1978; Tyler & Cook, 1984).

Previous and current results and conclusions, however, are not necessarily inconsistent. First and foremost, there is no doubt that because of fierce competition in the media marketplace, newsworthy items are screened for their potential to captivate the audience. Thus, the media tend to amplify some aspects of reality while scaling down others (Meyer, 1990). We can observe traces of this dynamic also in our data. Consistent with previous analyses, we found distortions in media coverage, such as an overemphasis of news coverage of homicide, accidents, and disasters. For instance, although only the 15th most frequent cause of death in the assorted set, homicide was the 4th most frequently reported cause of death in the media (see Table 4). This may have contributed to the overestimation of homicide, which was estimated to be the 8th most frequent cause of death. Further, although asthma is more than twice as frequent as homicide, these two risks were estimated to be equally frequent. People could retrieve many cases of homicide from their virtual circle but not a single case of asthma (see Table 4). Notwithstanding these individual examples, across *all* risks, people's judgments were relatively sheltered from media biases. This observation is in line with recent analyses. For instance, Sjöberg and Engelberg (2010) observed that the influence of media exposure on risk perception, for instance, in terms of movies and news that reiterate what is already known (e.g., the spike in media coverage at the 10-year anniversary of the Chernobyl accident), is, at best, small and short-lived.

Two Limitations

Both of our studies involved a predominately female student sample. Therefore, we currently do not know to what extent our observations will generalize to a public consisting of citizens with diverse educational backgrounds, various ages, and/or a balanced gender distribution. A second possible limitation is that we recorded naturally occurring instance knowledge and affective responses, but did not manipulate them experimentally. To ascertain the causal influence of availability and affect on risk judgments, one could, for instance, attempt to prime instance knowledge or heighten (or lower) participants' dread of specific risks (e.g., by highlighting a risk's catastrophic potential), and then record how these experimental manipulations shape people's judgments of the risks.

Practical and Policy Implications

Our findings have several implications for public policy. One concerns the question of how well people are calibrated to risks in the environment. As we have seen, the story is not simply that people's calibration is either poor or good. Instead, accuracy depends on how risk perception is measured (cf. Fischhoff & MacGregor, 1983). The number of recalled instances in one's social network is a good (i.e., ecologically valid) cue for inferring objective mortality (see Table 3). Consequently, asking people to judge risks in terms of their frequencies appears to trigger a heuristic that exploits this cue (cf. Hertwig et al., 2005), resulting in relatively good calibration. By contrast, asking people for perceived risk or VSL judgments also invites reliance on affective information, which is not an invariably good cue (see Table 3).

Of course, calibration to objective frequencies is only one measure of the public's assessment of risks. People's concerns and fears, for instance, regarding a risk's disaster potential are also important and tangible dimensions that can strongly affect legislative agendas (Slovic, 2000, p. 390). Our results suggest that probing such consequential fears requires using measures of risk that go beyond perceived mortality frequencies. We found that measures such as perceived risk and VSL judgments are more likely to trigger affective components associated with risks. Therefore, depending on policy makers' and risk analysts' objectives—for instance, helping to calibrate the public to objective mortality risks, thus attenuating the risk of misallocating public resources (Sunstein, 2002), or forecasting how the public will respond to technological risks and risk regulations—they should take different roads to citizens' intuitions about risks.

Our results also have several practical implications by suggesting an important and neglected dimension in the design of public health campaigns. A common approach to calibrate people better to the risks in the environment and to change their behavior is to provide them with aggregate frequency statistics about the occurrence of risks, despite the fact that these efforts are often met with limited success (cf. Slovic, 1986). An alternative and potentially more effective way to educate the public about a specific risk could be to encourage people to trust their own experience concerning the risk garnered from their proximate social network. If the risk claimed fewer victims in their own social circle than another risk, then—on average—this asymmetry is a rough but good proxy for the risks' relative population frequencies. For instance, in order to make people aware of obesity-associated risks, campaigns could start with people's direct experiences and make them cognizant of instances of obesity-related diseases in their social ecology. Of course, there are limits to this approach: Direct experience is most likely available for common risks and lacking for rare risks. Education campaigns, however, often target common risks. Admittedly, people's memories of risk occurrences in their social environment are likely not devoid of distortions (e.g., due to primacy, recency, or sampling error). Moreover, such memories of risk occurrences may be rich in affect. Despite these confounding factors, however, our analyses show that the number of instances recalled from people's social circles is, by and large, a quite valid indicator of the risks' mortality rates in the population.

Still another practical implication of our results for risk communication concerns the effectiveness of fear appeals (i.e., health-risk messages that arouse fear). Such appeals are often designed to amplify the amount of dread associated with a risk, and are used, for instance, to prevent drug consumption (for an overview, see Witte & Allen, 2000). Our results suggest conditions under which fear appeals are more or less likely to affect people's risk perception. For instance, the designers of fear appeals should consider how likely it is that people have direct experience with the risk in question. When people lack direct experience, fear appeals find an experiential tabula rasa on which their impact can unfold more easily. When people possess direct experience, the question is whether it concurs with the fear appeal (thus possibly amplifying its impact) or contradicts it (thus possibly rendering it ineffective). Finally, conclusions about the impact of fear appeals are likely to differ depending on how one defines and measures the target audience's risk perception (e.g., morality rates vs. VSL judgments).

Conclusions

The staggering rises in the cost of health care and the public-debt crises in many industrialized countries is increasing the pressure on societies to efficiently allocate their limited resources, so that the big problems and the big risks can be managed better. As Sullivan et al. (2011) highlighted, this process of allocating dwindling resources requires the input of important stakeholders and the public in general, in the health-care domain and beyond. Public participation is likely to increase the chance that agreed-upon priorities will be perceived to be legitimate. In order to veridically assess the public's priorities, it is imperative that we fully understand how the public judges risks and how different measures of risk perception may generate divergent subjective mappings of risk. By illuminating the differential roles of availability and affect in different facets of people's perception of risk, as well as the interplay between availability and affect, our findings contribute to a better understanding of the cognitive underpinnings of subjective risk. Our observations show that the accuracy of people's risk perception depends partly on how calibration is measured; they inspire an intuitive way to calibrate people's concerns about risks (e.g., encourage them to consult their direct experiences) and may aid the design of more effective health campaigns targeting the affective dimension of risks.

References

Betsch, T., & Pohl, D. (2002). Tversky and Kahneman's availability approach to frequency judgement: A critical analysis. In P. Sedlmeier & T. Betsch (Eds.), *Etc.: Frequency processing and cognition* (pp. 109–119). Oxford, UK: Oxford University Press.

Breckenridge, J. N., & Zimbardo, P. G. (2006). The strategy of terrorism and the psychology of mass-mediated fear. In B. Bongar, L. Beutler, L. M. Brown, J. N. Breckenridge, & P. G. Zimbardo (Eds.), *Psychology and terrorism.* (pp. 116–133). New York, NY: Oxford University Press.

Brown, N. R., & Siegler, R. S. (1993). Metrics and mappings: A framework for understanding real-world quantitative estimation. *Psychological Review, 100,* 511–534. doi:10.1037/0033-295X.100.3.511

Brunswik, E. (1952). The conceptual framework of psychology. *International Encyclopedia of Unified Science*. Chicago, IL: University of Chicago Press.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review, 62,* 193–217. doi:10.1037/h0047470

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences.* Hillsdale, NJ: Erlbaum.

Combs, B., & Slovic, P. (1979). Causes of death: Biased newspaper coverage and biased judgments. *Journalism Quarterly, 56, 837–849.* doi:10.1177/107769907905600420

Dhami, M. K., Hertwig, R., & Hoffrage, U. (2004). The role of representative design in an ecological approach to cognition. *Psychological Bulletin, 130,* 959–988. doi:10.1037/0033-2909.130.6.959

Eisenman, R. (1993). Belief that drug usage in the United States is increasing when it is really decreasing: An example of the availability heuristic. *Bulletin of the Psychonomic Society, 31,* 249–252.

Ekman, P., & Friesen, W. V. (1986). A new pan-cultural facial expression of emotion. *Motivation and Emotion, 10,* 159–168. doi:10.1007/BF00992253

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist, 49,* 709–724. doi:10.1037/0003-066X.49.8.709

Finucane, M. L., Alhakami, A., Slovic, P., & Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making, 13,* 1–17. doi:10.1002/(SICI)1099-0771(200001/03)13:1<1::AID-BDM333>3.0.CO;2-S

Fischhoff, B., Gonzalez, R. M., Lerner, J. S., & Small, D. A. (2005). Evolving judgments of terror risks: Foresight, hindsight, and emotion. *Journal of Experimental Psychology: Applied, 11,* 124–139. doi:10.1037/1076-898X.11.2.124

Fischhoff, B., & MacGregor, D. (1983). Judged lethality: How much people seem to know depends upon how they are asked. *Risk Analysis, 3,* 229–236. doi:10.1111/j.1539-6924.1983.tb01391.x

Fischhoff, B., Slovic, P., Lichtenstein, S., Read, S., & Combs, B. (1978). How safe is safe enough? A psychometric study of attitudes towards technological risks and benefits. *Policy Sciences, 9,* 127–152. doi:10.1007/BF00143739

Folkes, V. S. (1988). The availability heuristic and perceived risk. *Journal of Consumer Research, 15,* 13–23. doi:10.1086/209141

Gigerenzer, G., & Richter, H. R. (1990). Context effects and their interaction with development: Area judgments. *Cognitive Development, 5,* 235–264. doi:10.1016/0885-2014(90)90017-N

Hertwig, R., Pachur, T., & Kurzenhäuser, S. (2005). Judgments of risk frequencies: Tests of possible cognitive mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 621–642. doi:10.1037/0278-7393.31.4.621

Hills, T. T., & Pachur, T. (2012). Dynamic search and working memory in social recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 218–228. doi:10.1037/a0025161

Jacoby, L. L., Kelley, C. M., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology, 56,* 326–338. doi:10.1037/0022-3514.56.3.326

Johnson, E. J., & Tversky, A. (1983). Affect, generalization, and the perception of risk. *Journal of Personality and Social Psychology, 45,* 20–31. doi:10.1037/0022-3514.45.1.20

Johnson, E. J., & Tversky, A. (1984). Representations of perceptions of risk. *Journal of Experimental Psychology: General, 113,* 55–70. doi:10.1037/0096-3445.113.1.55

Jones, S. K., Jones, K. T., & Frisch, D. (1995). Biases of probability assessment: A comparison of frequency and single-case judgments. *Organizational Behavior and Human Decision Processes, 61,* 109–122. doi:10.1006/obhd.1995.1009

Kahneman, D. (2011). *Thinking fast and slow.* New York, NY: Farrar, Strauss, Giroux.

Keller, C., Siegrist, M., & Gutscher, H. (2006). The role of the affect and availability heuristics in risk communication. *Risk Analysis, 26,* 631–639. doi:10.1111/j.1539-6924.2006.00773.x

Länder vernichten Millionen Dosen H1N1-Impfstoff. (2011, November). *Spiegel Online*. Retrieved from http://www.spiegel.de/wissenschaft/medizin/0,1518,799993,00.html

Lerner, J. S., Gonzalez, R., Small, D. A., & Fischhoff, B. (2003). Effects of fear and anger on perceived risks of terrorism: A national field experiment. *Psychological Science, 14,* 144–150. doi:10.1111/1467-9280.01433

Lerner, J. S., & Keltner, D. (2000). Beyond valence: Toward a model of emotion-specific influences on judgment and choice. *Cognition and Emotion, 14,* 473–493. doi:10.1080/026999300402763

Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M., & Combs, B. (1978). Judged frequency of lethal events. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 551–578. doi:10.1037/0278-7393.4.6.551

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, E. S. (2001). Risk as feelings. *Psychological Bulletin, 127,* 267–286. doi:10.1037/0033-2909.127.2.267

Lorch, R. F., Jr., & Myers, J. L. (1990). Regression analyses of repeated-measures data in cognitive research: A comparison of three different methods. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 149–157. doi:10.1037/0278-7393.16.1.149

Meyer, P. (1990). News-media responsiveness to public health. In C. Atkin & L. Wallack (Eds.), *Mass communication and public health* (pp. 52–59). Newbury Park, CA: Sage.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81–97. doi:10.1037/h0043158

Oppenheimer, D. M. (2004). Spontaneous discounting of availability in frequency judgment tasks. *Psychological Science, 15,* 100–105. doi:10.1111/j.0963-7214.2004.01502005.x

Pachur, T. (2010). Recognition-based inference: When is less more in the real world? *Psychonomic Bulletin & Review, 17,* 589–598. doi:10.3758/PBR.17.4.589

Pachur, T. (2011). The limited value of precise tests of the recognition heuristic. *Judgment and Decision Making, 6,* 413–422.

Pachur, T., & Galesic, M. (in press). Strategy selection in risky choice: The impact of numeracy, affect, and cross-cultural differences. *Journal of Behavioral Decision Making.* doi: 10.1002/bdm.1757

Pachur, T., & Hertwig, R. (2006). On the psychology of the recognition heuristic: Retrieval primacy as a key determinant of its use. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 983–1002. doi:10.1037/0278-7393.32.5.983

Pachur, T., Hertwig, R., & Rieskamp, J. (in press). The mind as an intuitive pollster: Frugal search in social spaces. In R. Hertwig, U. Hoffrage, and the ABC Research Group (Eds.), *Simple heuristics in a social world.* New York, NY: Oxford University Press.

Park, J. S., & Grow, J. M. (2008). The social reality of depression: DTC advertising of antidepressants and perceptions of the prevalence and lifetime risk of depression. *Journal of Business Ethics, 79,* 379–393. doi:10.1007/s10551-007-9403-7

Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology, 61,* 195–202. doi:10.1037/0022-3514.61.2.195

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology, 45,* 513–523. doi:10.1037/0022-3514.45.3.513

Schweizerische Eidgenossenschaft, Bundesamt für Statistik. (2007). *Todesursachenstatistik: Ursachen der Sterblichkeit 2003 und 2004.* Neuchâtel, Switzerland: Author.

The shaming of America. (2005, Sept 8). *The Economist.* Retrieved from http://www.economist.com/node/4370617

Sjöberg, L. (2007). Emotions and risk perception. *Risk Management, 9,* 222–237. doi:10.1057/palgrave.rm.8250038

Sjöberg, L., & Engelberg, E. (2010). Risk perception and movies: A study of availability as a factor in risk perception. *Risk Analysis, 30,* 95–106. doi:10.1111/j.1539-6924.2009.01335.x

Slovic, P. (1986). Informing and educating the public about risk. *Risk Analysis, 6,* 403–415. doi:10.1111/j.1539-6924.1986.tb00953.x

Slovic, P. (1987). Perception of risk. *Science, 236,* 280–285. doi:10.1126/science.3563507

Slovic, P. (2000). *The perception of risk.* London, UK: Earthscan.

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

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1979). Rating the risks. *Environment, 21,* 14–39. doi:10.1080/00139157.1979.9933091

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1980). Facts and fears: Understanding perceived risk. In R. Schwing & W. A. Albers, Jr. (Eds.), *Societal risk assessment: How safe is safe enough?* (pp. 181–214). New York, NY: Plenum.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1985). Characterizing perceived risk. In R. W. Kates, C. Hohenemser, & J. X. Kasperson (Eds.), *Perilous progress: Managing the hazards of technology.* (pp. 91–125). Boulder, CO: Westview.

Slovic, P., Monahan, J., & MacGregor, D. G. (2000). Violence risk assessment and risk communication: The effects of using actual cases, providing instruction, and employing probability versus frequency formats. *Law and Human Behavior, 24,* 271–296. doi:10.1023/A:1005595519944

Slovic, P., & Peters, E. (2006). Risk perception and affect. *Current Directions in Psychological Science, 15,* 322–325. doi:10.1111/j.1467-8721.2006.00461.x

Small, D., Loewenstein, G., & Slovic, P. (2007). Sympathy and callousness: The impact of deliberative thought on donations to identifiable and statistical victims. *Organizational Behavior and Human Decision Processes, 102,* 143–153. doi:10.1016/j.obhdp.2006.01.005

Stewart, N., Chater, N., & Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology, 53,* 1–26. doi:10.1016/j.cogpsych.2005.10.003

Sullivan, R., Peppercorn, J., Sikora, K., Zalcberg, J., Meropol, N. J., Amir, E., . . . Aapro, M. (2011). Delivering affordable cancer care in high-income countries. *Lancet Oncology, 12,* 933–980. doi:10.1016/S1470-2045(11)70141-3

Sunstein, C. R. (2002). *Risk and reason: Safety, law, and the environment*. Cambridge, UK: Cambridge University Press.

Taylor, S. E., & Thompson, S. C. (1982). Stalking the elusive “vividness” effect. *Psychological Review, 89,* 155–181. doi:10.1037/0033-295X.89.2.155

Tengs, T. O., Adams, M. E., Pliskin, J. S., Safran, D. G., Siegel, J. E., Weinstein, M. C., & Graham, J. D. (1995). Five hundred life-saving interventions and their cost effectiveness. *Risk Analysis, 15,* 369–390. doi:10.1111/j.1539-6924.1995.tb00330.x

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review, 34,* 273–286. doi:10.1037/h0070288

Tsai, C. I., & Thomas, M. (2011). When does feeling of fluency matter? How abstract and concrete thinking influence fluency effects. *Psychological Science, 22,* 348–354. doi:10.1177/0956797611398494

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5,* 207–232. doi:10.1016/0010-0285(73)90033-9

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124–1131. doi:10.1126/science.185.4157.1124

Tyler, T. R., & Cook, F. L. (1984). The mass media and judgments of risk: Distinguishing impact on personal and societal level judgments. *Journal of Personality and Social Psychology, 47,* 693–708. doi:10.1037/0022-3514.47.4.693

Watkins, M. J., & LeCompte, D. (1991). Inadequacy of recall as a basis for frequency knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 1161–1176. doi:10.1037/0278-7393.17.6.1161

Witte, K., & Allen, M. (2000). A meta-analysis of fear appeals: Implications for effective public health campaigns. *Health Education & Behavior, 27,* 591–615. doi:10.1177/109019810002700506

Zanuck, R. D. (Producer), Brown, D. (Producer), & Spielberg, S. (Director). (1975). *Jaws* [Motion picture]. United States: Universal Pictures.

Appendix A

How Do People Judge Risks When Experienced Instances Do Not Discriminate?

Study 1

As reported, availability-by-recall does not discriminate between risks as frequently as the affect heuristic. Therefore, one could ask whether the latter may step in to help when the former does not render possible a decision. Specifically, how well does the affect heuristic predict relative risk judgments when people’s knowledge of specific instances does not help them? In the subsets of comparisons where instances did not discriminate, the affect heuristic made predictions for, on average, 61.3% (SD = 22.0; frequency judgments) and 43.8% (SD = 37.4; VSL judgments) of the cases. Out of those, the affect heuristic correctly predicted 61.1% (SD = 16.7) for the frequency judgments and 64.9% (SD = 22.9) for the VSL judgments, substantially higher than chance (i.e., 50%; $t > 2.98$, $p < .007$). In other words, rather than guessing when instance knowledge does not discriminate between two risks, people seem to turn to affective information.

Study 2

Table A1 shows the predictive accuracy of availability-by-recall$_{\mathrm{TEX}}$ and the two instantiations of the affect heuristic for those cases in which direct experience fails to discriminate. As can be seen, in those cases, availability-by-recall$_{\mathrm{TEX}}$ does not do better than chance; both instantiations of the affect heuristic, however, do. As in Study 1, we thus find evidence consistent with the possibility that people resort to affective information when direct experience fails to discriminate between risks.

(*Appendices continue*)

Table A1

Mean Percentage of Items (from those in which Participants' Instance Knowledge Did not Discriminate Between the Risks) That Could Be Modeled With the Two Instantiations of the Affect Heuristic and Availability-by-recall$_{TEX}$ in Study 2; and the Percentage of Correct Predictions for the Three Models, Separately for the Frequency-Choice Task, the Value-of-a-Statistical-Life (VSL) Choices, and the Perceived-Risk Choice Task

	Choice	Affect heuristic (dread score)	Affect heuristic (dread item)	Availability-by-recall$_{TEX}$ (direct and indirect experience)
Percentage of items modeled	Frequency	96.4% (2.3)	68.0% (18.2)	50.9% (26.4)
	VSL	92.7% (20.4)	68.6% (23.1)	47.3% (27.3)
	Perceived risk	93.9% (15.5)	64.2% (21.9)	46.0% (28.5)
Performance	Frequency	58.3%* (8.5)	64.8%* (10.5)	49.5% (12.7)
	VSL	59.5%* (13.7)	65.5%* (15.2)	50.8% (18.5)
	Perceived risk	62.7%* (7.1)	69.7%* (11.9)	52.5% (15.4)

Note. Standard deviations are shown in parentheses.
* Performance is significantly better than chance ($p < .05$).

Appendix B

Correlation and Regression Analyses for Individual Participants

In addition to the correlation and regression analyses on the aggregate level, we also conducted these analyses on the individual level. As regards the regression analyses, we conducted a random coefficient analysis (Lorch & Myers, 1990), which is a multilevel regression technique in which the magnitude of an experimental effect is first estimated within each participant, and then these within-subjects effects are tested against the null hypothesis of no effect.

Study 1

Although, due to the larger amount of unsystematic variance in individual data the correlations were lower than for the aggregate analyses, the correlations on the individual level replicated the results on the aggregate level: Estimated frequencies were, on average (across participants), more strongly correlated to the availability of instances than to the amount of dread a risk evokes, mean *r*s = 0.32 ($SD = 0.18$) versus 0.16 ($SD = 0.44$), $t(29) = 1.82$, $p = .08$, $d = .48$ (paired-samples *t* test based on *z*-transformed individual correlations); and there was a trend in the opposite direction for VSL judgments, with mean correlations of 0.15 ($SD = 0.22$) to availability and 0.29 ($SD = 0.40$) to affect, $t(18) = -1.41$, $p = .18$, $d = -.43$. As on the aggregate level, it was also the case on the individual level that availability and frequency estimates were correlated even when controlling for actual frequency, individual *r*s = .26, $t(29) = 7.35$, $p = .001$, $d = 1.90$ (one-sample *t* test against zero).

As Table B1 shows, the mean (across participants) standardized regression coefficients were higher for availability than for affect in predicting absolute frequency estimates (although again, the amount of explained variance was lower than for the aggregate analyses). In predicting VSL judgments, by contrast, the relative contributions of availability and affect reversed, and hence affect had the larger (mean) regression coefficient.

Study 2

Table B2 shows, using the individual-level zero-order correlations, how the frequency estimates and the VSL judgments were related to the measures of availability and affect. Overall, the same pattern was obtained as on the aggregate level: The frequency estimates were correlated with direct experience (i.e., the number of recalled instances), but only weakly to indirect experience (i.e., the number of instances recalled from the media). Both the frequency estimates and the VSL judgments were about equally strongly related to the measures of affect (as measured by the single dread item) as to direct experience. As in Study 1, the correlation between frequency estimates and direct experience also held when controlling for actual frequency, mean individual *r*s = .19, $t(29) = 9.44$, $p = .001$, $d = 1.47$ (one-sample *t* test against zero).

Table B1 shows that, although due to more unsystematic variance in individual data the differences in the regression coefficients for affect and availability were less pronounced, the same qualitative patterns were obtained as on the aggregate level. Specifically, for the frequency estimates, direct experience had the largest regression weight, whereas for VSL judgments, the dread item had the largest weight. Overall, indirect experience had only a small—and negative—impact on the risk assessments.

(*Appendices continue*)

Table B1
Results of the Individual-Participant Regression Analyses

	Study 1		Study 2	
Predictor	Frequency estimates (log)	VSL (log)	Frequency estimates (log)	VSL (log)
Dread score	**0.16** (0.43)	**0.27** (0.39)	**0.22** (0.25)	**0.16** (0.26)
Dread item	—	—	**0.31** (0.26)	**0.27** (0.28)
Direct experience (log)	**0.26** (0.18)	**0.10** (0.18)	**0.32** (0.19)	**0.20** (0.19)
Indirect experience (log)	—	—	**0.07** (0.20)	−0.02 (0.20)
R^2_{adj}	.18	.15	.47	.31

Note. For the analyses, each participant's log-transformed frequency estimates and value-of-a-statistical-life (VSL) judgments for the different risks were regressed on the different measures of affect (dread score and dread item) and availability (number of instances in social network = direct experience, number of instances in media = indirect experience). Shown are the means of the individual-participant standardized regression coefficients (*SD*s in parentheses) as well as the mean adjusted R^2. Regression coefficients significantly larger than zero ($p < .05$, based on a one-sample *t* test against zero) are in bold.

Table B2
Mean Individual-Participant Intercorrelations Between the Different Measures of Availability and Affect, Their Correlation to the Actual Frequencies of the Risks and to the Frequency Estimates and Value-of-a-Statistical-Life (VSL) Judgments in Study 2

	Availability		Affect	
	Direct experience (log)	Indirect experience (log)	Dread score	Dread item
Direct experience (log)	—	**.27**	**.31**	**.37**
Indirect experience (log)		—	**.35**	**.26**
Dread score			—	**.57**
Actual frequency (log)	**.53**	**.24**	**.50**	**.53**
Estimated frequency (log)	**.50**	**.26**	**.46**	**.53**
VSL (log)	**.34**	**.11**	**.33**	**.40**

Note. Significant correlations ($p = .05$; two-tailed) are in bold.

Received July 5, 2011
Revision received March 6, 2012
Accepted March 22, 2012 ■