THINKING & REASONING, 2017
https://doi.org/10.1080/13546783.2017.1307785

Risk it? Direct and collateral impacts of peers' verbal expressions about hazard likelihoods*

Paul D. Windschitl[a], Andrew R. Smith[b], Aaron M. Scherer[c] and Jerry Suls[d]

[a]Department of Psychological and Brain Sciences, University of Iowa, Iowa City, Iowa, USA; [b]Department of Psychology, Appalachian State University, Boone, North Carolina, USA; [c]Department of Internal Medicine, University of Iowa, Iowa City, Iowa, USA; [d]Behavioral Research Program, National Cancer Institute, Bathesda, Maryland, USA

ABSTRACT
When people encounter potential hazards, their expectations and behaviours can be shaped by a variety of factors including other people's expressions of verbal likelihood (e.g., *unlikely to harm*). What is the impact of such expressions when a person also has numeric likelihood estimates from the same source(s)? Two studies used a new task involving an abstract virtual environment in which people learned about and reacted to novel hazards. Verbal expressions attributed to peers influenced participants' behaviour toward hazards even when numeric estimates were also available. Namely, verbal expressions suggesting that the likelihood of harm from a hazard is low (vs. higher) yielded more risk taking with respect to said hazard. There were also inverse *collateral effects*, whereby participants' behaviour and estimates regarding another hazard in the same context were affected in the opposite direction. These effects may be based on directionality and relativity cues inferred from verbal likelihood expressions.

ARTICLE HISTORY Received 28 July 2016; Accepted 5 March 2017

KEYWORDS Risk perception; decision making; verbal probabilities; risk communication

Introduction

People often share opinions with others about the frequencies or likelihood of important events. Sometimes, these opinions have two parts: one part in which the speaker uses verbal phases to capture their impression of frequency or likelihood, and another part in which the speaker numerically estimates the relevant percentage. Consider some examples. "I'd say your chances are good—around 70%". "Sheri rarely misses these problems; she catches 85% of them". "Only 20% of male customers want the color white; it's unlikely to see ones that do". For these situations, the percentage information

---

CONTACT Paul D. Windschitl ✉ paul-windschitl@uiowa.edu

*The statements and opinions stated in this manuscript do not necessarily represent the opinions of the National Cancer Institute.

may seem to be the clear and complete information that a listener needs for assessing the relevant, objective frequency or likelihood. The percentage is interpretable on a common-rule scale from 0 to 100, on which a given number should mean exactly the same thing across people and contexts (Biernat, Manis, & Nelson, 1991). Verbal phrases, however, do not enjoy this simplicity (Beyth-Marom, 1982; Brun & Teigen, 1988; Budescu & Wallsten, 1995; Weber & Hilton, 1990). This raises the question of whether people put much stock in verbal phrases when exposed to dual mode communications that include both numeric estimates and verbal descriptors. This paper examines the influence of such verbal phrases on listeners' behaviour and beliefs regarding not only the event about which the speaker was referring, but also other events within the same context. In other words, this paper examines the direct and collateral effects of verbal phrases in these dual mode communications. These issues were examined within a novel paradigm in which participants received information about potential hazards in an abstract virtual environment and had to make decisions about how to behave with respect to those hazards.

The potential influence of verbal expressions

For communicating proportions and probabilities, a numeric format (e.g., "there is a 15% chance") has important, desirable characteristics. The 0–100 scale is a common-rule scale on which values should have a fixed meaning (Biernat et al., 1991). Also, values can be derived from event frequencies, and their accuracy can often be measured against objective data. Perhaps aware of some of these desirable characteristics, people often express a preference for receiving probabilistic information in numerical form (Erev & Cohen, 1990; Olson & Budescu, 1997; Shaw & Dear, 1990; see also Gurmankin, Baron, & Armstrong, 2004a, 2004b). For numerate individuals receiving input from a peer about an observed proportion or an inferred probability, it might be difficult to imagine a better format in which to receive the information. The most obvious alternative format—namely, receiving verbal phrases that describe the proportion or probability—has potential pitfalls. Those expressions (e.g., "there is a small chance") can be vague and flexibly interpreted. Translations of verbal expressions into numerical forms reveal substantial variability across persons, cultures, contexts, and event types (e.g., Beyth-Marom, 1982; Brun & Teigen, 1988; Budescu & Wallsten, 1995; Budescu, Broomell, & Por, 2009; Budescu, Por, Broomell, & Smithson, 2014; Fischer & Jungermann, 1996; Harris & Corner, 2011; Harris, Corner, Xu, & Du, 2013; Piercey, 2009; Wallsten, Budescu, Rapoport, Zwick, & Forsyth, 1986; Weber, 1994; Weber & Hilton, 1999).

Nonetheless, for this project we expected that even when people receive precise numeric estimates from peers, verbal information from those peers will exert a significant impact on the receivers. There are many interrelated reasons for this expectation. First, large segments of the population struggle with basic numeracy (Gigerenzer, Hertwig, Van Den Broek, Fasolo, &

Katsikopoulos, 2005; Lipkus, Samsa, & Rimer, 2001; Peters et al., 2006). It is, therefore, possible that these people prioritise verbal expressions over estimates in the form of numbers.

Second, even for relatively simple statistics for which numerical confusion is typically not an issue, a statistic can seem rather pallid and might have surprisingly little influence relative to anecdotal information (Borgida & Nisbett, 1977). If anecdotes can swamp the influence of more reliable numeric information, other types of information, such as verbal expressions, might show a similar effect.

Third, a statistic can be of limited meaning or evaluability (Hsee, 1996; Windschitl, Martin, & Flugstad, 2002; Zikmund-Fisher, 2013; Zikmund-Fisher, Fagerlin, & Ubel, 2004). A numeric risk estimate, per se, communicates nothing about interpretation; it does not provide information about whether the source perceives the risk to be high, moderate, or low (Teigen & Brun, 1995, 2000; Windschitl et al., 2002). In contrast, verbal expressions tend to be directional; they reveal something about how the source of the expression views the magnitude of likelihood and perhaps even subtle opinions about what should be done (Honda & Yamagishi, 2006; Juanchich, Sirota, Karelitz, & Villejoubert, 2013; Moxey & Sanford, 2000; Sanford & Moxey, 2003; Teigen & Brun, 2000, 2003a). Consider the terms "extremely unlikely” and “not entirely impossible” They might be translated to similar numeric probabilities, but they seem to point in different directions. The former implies the listener should consider an event's non-occurrence, whereas the latter to consider its occurrence (see Teigen & Brun, 2003b, for a discussion of consequences of phrase directionality).

Finally, even if numeric forecasts are interpreted in a precise way at a deliberative level, people's intuitive interpretations may remain malleable (Windschitl & Weber, 1999). For example, although a 5% estimate of rain in London or in Madrid can be understood as meaning the same probability, the more intuitive interpretation of this 5% value can be influenced by associations people have regarding weather in those cities (see also Bilgin & Brenner, 2013; Flugstad & Windschitl, 2003). Related research has revealed this type of intuitive flexibility in other ways, such as showing that people's reactions to a numeric value can be influenced by frames (e.g., 15% chance this plan will succeed vs. 85% chance this plan will not succeed) or even by other numbers that are presented within the same context (Bigman, 2014; Fagerlin, Zikmund-Fisher, & Ubel, 2007; Klein, 1997; Levin, Schneider, & Gaeth, 1998; McKenzie & Nelson, 2003; Windschitl et al., 2002). This type of intuitive flexibility, or the separability of intuitive and deliberative representations of uncertainty, offers a way for verbal expressions to impact expectations even when numeric information is known and accepted.

In short, despite the strength of a numeric format for communicating likelihood, we expected that a peer's verbal expressions that are presented with a numeric estimate would have distinct influences on a person receiving that communication. The verbal expressions could influence intuitive representations of uncertainty that are conceptually different from participants’ numeric assessments of uncertainty. Assuming the intuitive representations influence behaviour, verbal expressions would ultimately also influence behaviour (e.g., in the context of our experiments they might influence whether people actually avoid a potential hazard). We were also interested in whether this influence would extend not only to how people behave regarding a focal hazard (about which the peer commented), but also to another hazard that could be encountered within the same context. We discuss this possibility next.

The potential collateral effects of verbal expressions

As mentioned earlier, verbal expressions of uncertainty tend to be directional (Honda & Yamagishi, 2006; Juanchich et al., 2013; Moxey & Sanford, 2000; Sanford & Moxey, 2003; Teigen & Brun, 2000, 2003a, 2003b). A peer's verbal expressions regarding the likelihood of a potential hazard causing harm could reveal something about not only the peer's estimate of the objective likelihood, but also a secondary evaluation of this level of likelihood and perhaps even his or her behavioural inclination in the situation. For example, if a peer says that a potential hazard is "extremely unlikely” to be a problem, this suggests a low objective likelihood and hints that the peer intuitively evaluates this level as low. Such intuitive evaluations, however, are rarely made in a vacuum. Evaluative judgments about virtually any target are impacted by points of reference, such as prior expectations or expectations derived from the category membership of the target (Hsee & Zhang, 2010; Huttenlocher & Higgins, 1971; Kahneman & Miller, 1986; Parducci, 1965; Sher & McKenzie, 2014; Vlaev, Chater, Stewart, & Brown, 2011). In other words, such evaluations tend to be relative judgments. We propose that listeners implicitly assume these are relative judgments, at least to some degree. Consequently, when a judgment context involves two salient hazards, a peer's verbal expression about one can influence a listener's beliefs and responses about the other. If a peer says that a potential hazard is "extremely unlikely” to be a problem, it may imply that the other potential hazard is more likely to be a problem than is the focal hazard. In short, verbal expressions about one hazard can have a collateral, inverse effect on interpretations about another hazard. In our studies, we test for this collateral impact by measuring behaviours and likelihood estimates regarding a hazard for which the participants did not receive information (i.e., a nonfocal hazard) in addition to a focal hazard.

Overview of tasks, studies, and predictions

We conducted two very similar studies with the same design, and both used a novel, computerised testing paradigm called the Decisions about Risk Task (DART). The DART is an abstract analogue of everyday situations in which people learn about potential hazards in multiple ways, and then decide how to behave—namely whether to take risks regarding those hazards. There is a learning phase and testing phase in DART. In the test phase (see Figure 1(a)) participants use the mouse to control a person icon and attempt to gain points by catching a meandering yellow target that reappears after each catch. Periodically, a potential hazard represented by a geometric shape appears (see Figure 1(b)), and after a short buffer period it attacks or not (see Figure 1(c) and 1(d)). During the buffer period, a participant must use experiences and information from the learning phase (described below) to inform a decision between two options: (1) a risky option in which they continue chasing targets and are open to attack, or (2) a safe option in which they move to the safe zone, where they are impervious to any attack but are kept from chasing targets for six seconds.

In the prior learning phase, participants witnessed trials in which two types of hazards—blue ovals and orange rectangles—attacked at the same rate (30%). In a 2 × 3 design, we manipulated the information participants received about one of the hazards (which we will call the focal hazard) at the very end of the learning phase. Participants were told that top-performing peers had estimated the focal hazard's numeric attack rate, and we varied

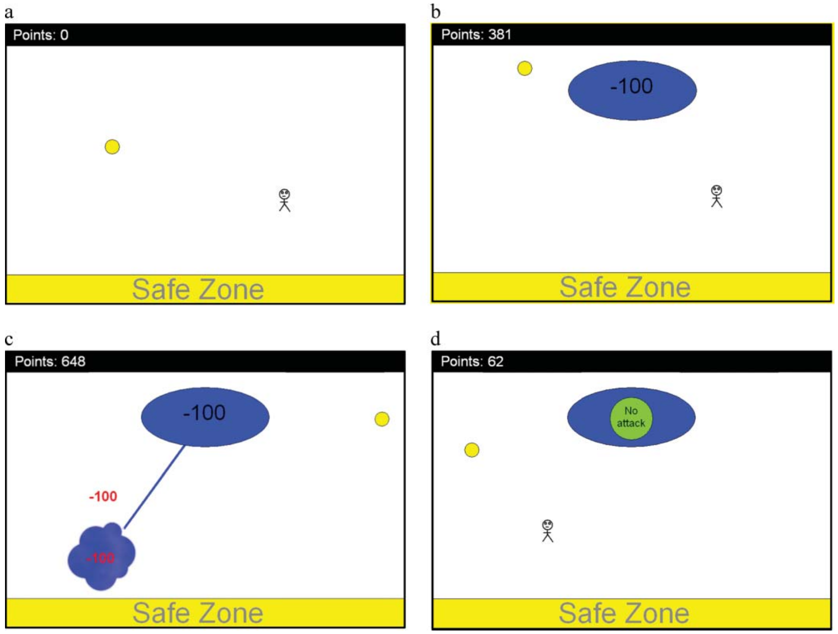

Figure 1. (a) Shows a screen shot from the testing phase of the DART. The dot is a randomly moving target, and participants gain points by chasing and "catching" the target with the person icon. Periodically, a potential hazard or threat, represented by a geometric shape, appears near the top center of the screen (see Figure 1(b)). The number in the shape, which was always −100 in the current studies, represents the penalty if successfully attacked by the hazard. After a 4-second buffer period, the potential hazard will either attack (see Figure 1(c)), which can result in the penalty, or it will harmlessly resolve in "no attack" (see Figure 1(d)).

whether the average estimate was reported as 20% or 40%. More important, we manipulated the verbal phrases that were said to come from those top-performing peers about the focal hazard: (1) expressions suggesting an infrequent attack rate, (2) no verbal expressions, or (3) expressions suggesting a frequent attack rate. Participants received no information attributed to peers about the nonfocal hazard.

The impact of the numeric manipulation is only of limited interest here, because the size of the manipulation was necessarily small (20% vs. 40%). We elected to manipulate that variable rather than hold it constant (e.g., at 30%), but we kept the manipulation small so that participants would view this information as plausible. Our primary interest was in participants' behaviour in the presence of focal and nonfocal hazards, as a function of the verbal information. In the presence of focal hazards, we expected that participants would go to the safe zone (i.e., make the safe choice) more often when verbal terms from peers suggested a frequent rather than infrequent attack rate, even though participants had, themselves, witnessed the same learning phase as the peers and were also given numeric information attributed to the peers. We were also able to test whether verbal phrases about the focal hazards had a collateral, inverse effect on behaviour toward nonfocal hazards.

In addition to measuring behaviour, we solicited verbal and numeric likelihood estimates from participants in the final trials of the test phase. On an exploratory basis, we also included a set of individual difference measures, which included two numeracy measures. Again, Studies 1 and 2 differed only slightly. The main difference between the studies concerned a variation in the delivery of the numeric probability information from peers.

Study 1

Method

Participants and design

The participants (*N* = 208) were students from introductory psychology courses at the University of Iowa. The design was a 2 (peers' numeric estimate: 20% or 40%) × 3 (peers' verbal estimate: absent, low, or high) between-subject design.[1] Blocking and counterbalancing factors were also included, as discussed below.

---

[1] At the start of data collection (Fall, 2010), there were two additional conditions included in our design that tested hypotheses quite different from those discussed here. Neither included any verbal information from peers. In one, participants received peers' numeric estimates about both hazards rather than just one. In the other, participants were given peers' numeric estimates about one hazard, but they were asked to guess the numeric attack rate before receiving it from peers. After collecting data from 91 participants for the three conditions reported here (and doing preliminary analyses), we decided to split the design into two separate studies, and we collected data from another 117 to arrive at our full sample of 208 in this study. This process was not optimal, but it should have little bearing on the interpretation of the study's results. The final sample size was based on a target of 200 participants (shaped by informal power estimates) coupled with variation due to lab logistics (time of semester, sign-up rates, availability of experimenters).

Procedure and task

The DART was introduced to participants as a game with a goal of scoring as many points as possible. The experimenter mentioned that if they scored "enough points", they would be allowed to pick a candy bar from a selection that was visible to the participants. Instructions included the following pieces of information about the test phase: (1) screenshots showing the basic elements, (2) points are earned by catching targets, (3) potential hazards called "threats" will occasionally appear and will either attack or not attack, (4) going to the safe zone protects from any attack but keeps one from collecting points for a period of six seconds, (5) staying outside the safe zone leaves one vulnerable to a 100 point loss from being attacked, and (6) there are two different threats. At this point, there was a short practice phase that mimicked the testing phase but involved different threats.

After the practice phase, participants were told that in order to perform well in the game it is useful to know about the attack tendencies of the threats, and that "the next phase of the study will help you learn these tendencies". They were then introduced to the learning or "observation" phase. In this phase, which involved the two hazards on which they would be tested, they witnessed serial presentations in which a hazard appeared onscreen and either attacked the person icon or displayed a no-attack sign. Each hazard (a blue oval and orange rectangle) appeared exactly 50 times and attacked on 30% of those occasions. To ensure participant attention, we had participants report which threat appeared on each trial and whether it attacked (both as dichotomous choices).

Following the observation phase, participants read: "Last semester, we had participants play the game you are playing. The information you will receive comes from the top three performers from last semester". Then participants received a verbal estimate (if applicable) and a numeric estimate. The estimates were always about the focal hazard; nothing was said about the nonfocal hazard. In cases where the focal threat was the blue oval (which was counterbalanced across participants), the verbal estimates were introduced with this text: "Previous participants were asked to comment on the blue oval. Here is what the top performers said". This was followed by three individually listed sentences: (1) "The blue oval [didn't seem to attack very often / seemed to attack quite often].", (2) "I [don't think / think] the blue oval attacks a lot.", and (3) "I'm not sure what to say, but it felt like the blue oval [rarely attacked / attacked a bunch of times].".

The numeric estimate was introduced with this text: "Previous participants were asked how often the shapes attacked after appearing. The top three performers from last semester varied slightly in their estimates, but showed generally good agreement. Below you can see the averages of their estimates for one of the shapes. The average attack rate estimate for the blue oval from the top performers last semester was [20%/40%]".

After some additional reminders, the test phase started. The participant's main task was to chase a yellow target that traveled in a meandering and unpredictable way. Once the target was caught, the participant earned 1, 5, or 10 points (randomly determined on each catch) and then resumed chasing the target when it reappeared in a new location. After a certain period of time, a potential hazard would appear at the top of the screen, which constituted the start of a trial. After a 4-second buffer, the hazard would either attack or not (again, see Figure 1(c) and 1(d) for illustrations). The computer recorded whether the participant made a risky behavioural choice (continued chasing targets) or a safe choice (moved to the safe zone) during the 4-second buffer for the given trial. Participants who went to the safe zone were impervious to any attack (losing 0 points even if an attack occurred, but missing out on opportunities to collect points while stuck in the safe zone for 6 seconds), whereas those outside the safe zone lost 100 points if the hazard did attack.[2]

Overall, there were 3 blocks of 20 trials. Within each block, half of the trials involved the focal hazard and half involved the nonfocal hazard. Both of the hazards continued to attack at the rate observed in the observation phase (i.e., 30%). Block 1 operated just as described in the above paragraph. Block 2 proceeded like Block 1 except that on 6 of the 20 trials, the action would freeze after a hazard appeared (but before it attacked or did not attack) and the participant would be asked to type a numeric estimate (0%–100%) for the likelihood that the threat would attack. Block 3 proceeded like Block 2, except that instead of making numeric estimates, the participants made verbal estimates on a 7-option scale ranging from *extremely unlikely* to *extremely likely*.

After the DART, participants completed an exit questionnaire on paper, which solicited a variety of information and responses (e.g., gender, age, frequencies of threats attacking during observation and test phases, beliefs about various aspects of the experiment). Participants also completed a numeracy scale (Lipkus et al., 2001), a subjective numeracy scale (Fagerlin et al., 2007), the Need for Cognition Scale (Cacioppo, Petty, & Kao, 1984), and a dispositional optimism scale (Scheier, Carver, & Bridges, 1994). All

---

[2] The precise expected value for the risky option is difficult to compute because it depends, in part, on how quickly participants implement their decisions (going to the safe zone early rather than late during the 4-second buffer sacrifices more point-gaining opportunities), how quickly they proceed out of the safe zone when released, and how generally efficient they are at chasing targets and gaining points. It is even possible that going to the safe zone occasionally serves as a rest, which might make people more efficient in immediately subsequent trials in which they stay out of the safe zone. Informal simulations suggest that a person who chases targets with moderate intensity would earn 50–60 points on most trails within 6 seconds, which is the length of time that one is precluded from catching targets when in the safe zone. The penalty of being attacked is 100 points. Given all this information, we could tentatively estimate that the expected value of the risky option is greater than 0 points for any hazard that attacks less than 55% of the time. The expected value of going the safe zone is 0 points.

participants received a large or small candy bar depending on their score. Finally, participants were debriefed.

Results

Preliminary analyses

Although participants in this study were attempting to maximise their overall score on the DART, this score was not of specific interest to us. Instead, our interest was in participants' behavioural decisions (and likelihood estimates) on focal and nonfocal trials, as a function of our manipulated variables. Nonetheless, as a preliminary matter, it is instructive to note that, given the specific DART parameters that we used in this study and given the general quickness with which participants pursued and caught targets in the task, participants tended to use the safe zones too often. For participants with at least average speed at catching targets, they should never use the safe zone. Even for hazards thought to attack at 40% (or perhaps as high as 55%), the expected value of going to the safe zone is less than the expected value of continuing to chase targets (see footnote 2 for more information about this estimation). In further support of the notion that participants went to the safe zone more than they should, there was a negative correlation between the percentage of trials participants went to the safe zone and their final score ($r = -.30$, $p < .001$).

Other preliminary analyses revealed no notable effects for gender, age, or focal-hazard counterbalancing on the primary dependent variables. Consequently, those factors are not discussed further.

Another preliminary point concerns the role of blocks. Participants' overall tendency to go to the safe zone decreased across blocks ($p < .001$). The average rates were 51.0%, 40.1%, and 36.5% for Blocks 1, 2, and 3, respectively. The drop occurred regardless of hazard type (focal or nonfocal), and it could be attributed to one or a mix of many processes that are not of particular interest here (e.g., learning, fatigue, the presence of likelihood questions in Blocks 2 and 3). Therefore, we do not discuss the blocking factor further, except to footnote the only cases—two—in which fuller analyses revealed that the blocking factor was involved in a significant interaction. All key findings below are the same whether or not block is included as a factor in analyses.

Decisions about the focal hazard

For each participant, we calculated the percentage of times he or she went to the safe zone within the trials in which a focal hazard appeared. We submitted these rates to a 2 (numeric estimate) × 3 (verbal estimate) analysis of variance (ANOVA). Figure 2 and Table 1 display the relevant means. The main effect of the numeric-estimate factor was relatively small and marginally significant,

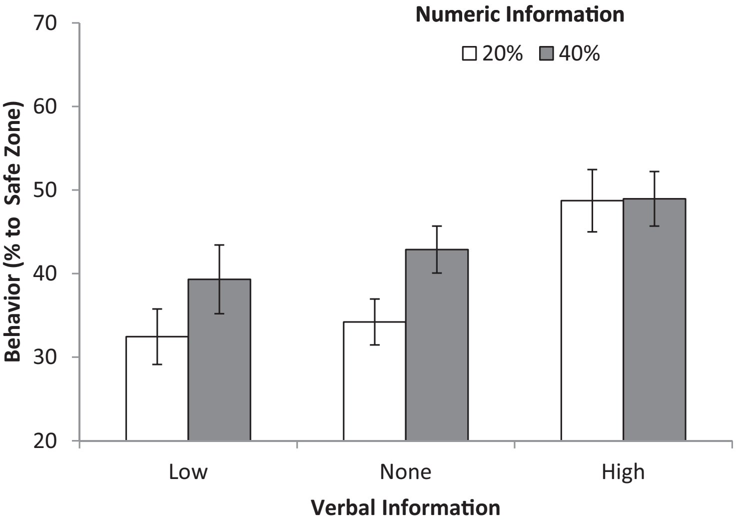

Figure 2. Participants' behavioural decisions (the mean percent of times they went to the safe zone) when faced with the focal hazard in Study 1, as a function of the numeric and verbal information they received from purported peers about the focal hazard. Error bars indicate ±1 standard errors.

$F(1, 202) = 3.68$, $p = .06$, $\eta_p^2 = 0.02$; participants were slightly more likely to go to the safe zone (43.7% vs. 38.5%) in response to a focal hazard when peers' estimates of that hazard's attack rate were said to be 40% rather than 20%. The fact that this effect is small may not be surprising because, as discussed above, the manipulation itself (40% vs. 20%) was kept within a narrow range so that participants would not reject the plausibility of this information attributed to peers.

More importantly, there was a significant main effect of the verbal-estimate factor, $F(2, 202) = 8.28$, $p < .001$, $\eta_p^2 = 0.08$. This effect reveals that, as expected, peers' verbal estimates influenced people's behaviour even though they also received peers' numeric estimates. Participants were most likely to go to the safe zone in the high-estimate condition and least likely in the low estimate condition ($p < .001$ for the difference). Other Tukey comparisons revealed that the no-estimate condition was different from the high-estimate condition ($p = .008$) but not the low estimate condition ($p = .67$). The interaction between the numeric and verbal factors was not significant, $F(2, 202) = 0.88$, $p = .42$, $\eta_p^2 = 0.01$. That is, the influence of the verbal information was similar regardless of whether the hazard was said to attack 20% or 40% of the time.

Decisions about the nonfocal hazard

We also examined how the peers' information about focal hazard impacted participants' behaviour toward the nonfocal hazard (see Figure 3 and Table 1).

Table 1. Descriptive statistics for behaviour, numeric estimates, and verbal estimates for both focal and nonfocal hazards in Study 1.

	Verbal condition							
Dependent variable	Low		None		High		Total	
Numeric condition	M	SD	M	SD	M	SD	M	SD
Focal, behaviour								
20%	32.5	19.4	34.2	16.0	48.7	21.8	38.5	20.4
40%	39.3	24.0	42.9	17.1	49.0	19.3	43.7	20.4
Total	35.9	21.9	38.7	17.0	48.8	20.4	41.2	20.5
Focal, NLE								
20%	34.4	16.3	28.2	12.9	38.2	17.1	33.6	15.9
40%	40.1	15.0	42.0	14.0	39.0	14.0	40.4	14.2
Total	37.2	15.8	35.4	15.1	38.6	15.5	37.0	15.4
Focal, VLE								
20%	3.6	0.9	3.5	1.2	4.4	0.9	3.8	1.1
40%	3.5	1.3	3.7	0.9	4.1	1.1	3.8	1.1
Total	3.6	1.1	3.6	1.1	4.3	1.0	3.8	1.1
Nonfocal, behaviour								
20%	48.8	18.5	39.7	13.7	40.1	18.6	42.9	17.5
40%	57.7	21.3	44.4	16.2	36.3	17.4	46.0	20.2
Total	53.3	20.3	42.2	15.1	38.2	18.0	44.5	18.9
Nonfocal, NLE								
20%	42.0	17.2	43.4	15.0	33.1	14.9	39.5	16.2
40%	49.1	14.5	44.9	14.7	37.1	16.0	43.7	15.7
Total	45.5	16.2	44.2	14.8	35.1	15.5	41.6	16.1
Nonfocal, VLE								
20%	3.9	0.9	3.9	1.0	3.6	0.7	3.8	0.9
40%	4.0	0.8	3.7	0.9	3.7	1.1	3.8	0.9
Total	4.0	0.9	3.8	0.9	3.7	0.9	3.8	0.9

Note: Behaviour = % of times going to safe zone when hazard was present. NLE = Numeric Likelihood Estimate (0%–100%). VLE = Verbal Likelihood Estimate (score from 1 to 7). Condition information refers to what participants were told that peers had said about the focal hazard (never the nonfocal hazard).

A 2 (numeric estimate) × 3 (verbal estimate) ANOVA on safety-zone rates in nonfocal trials revealed that, not surprisingly, there was not a main effect of numeric information, $F(1, 202) = 1.77$, $p = .19$, $\eta_p^2 = 0.01$. Critically, however, the verbal estimate main effect was significant, $F(2, 202) = 13.37$, $p < .001$, $\eta_p^2 = 0.12$. As discussed earlier, the direction of this effect was inverse. Participants were more likely to go to the safe zone for the nonfocal hazard when the focal hazard had been described with phrases suggesting a low rather than high likelihood of attack ($p < .001$). Other Tukey comparisons revealed that the mean for the no-estimate condition, which was intermediate relative to the other two conditions, was significantly different from the low-estimate condition ($p < .001$) but not the high estimate condition ($p = .38$). The interaction was not significant, $F(2, 202) = 2.30$, $p = .10$, $\eta_p^2 = .02$.[3]

---

[3] When block is included for a 2 (numeric estimate) × 3 (verbal estimate) × 3 (block) ANOVA on decisions from nonfocal trials, the block factor is involved in a significant two-way interaction with verbal estimate ($p = .02$) and a significant three-way interaction ($p = .04$). However, these unexpected interactions do not seem especially important, and they do not replicate in Study 2.

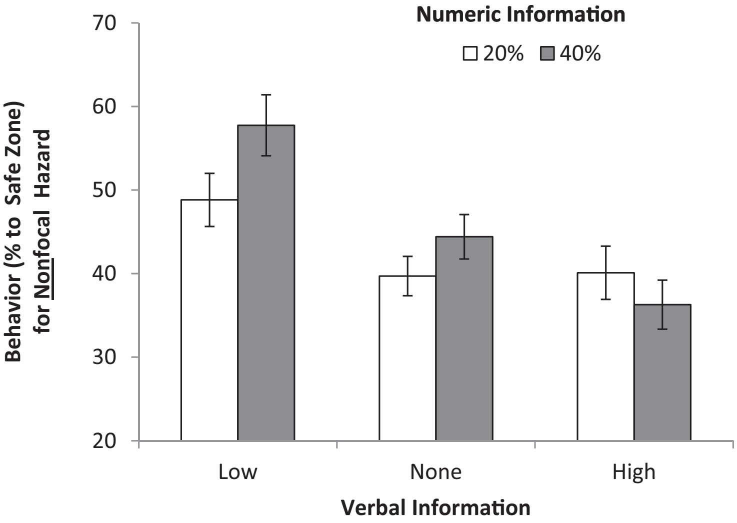

Figure 3. Participants' behavioural decisions (the mean per cent of times they went to the safe zone) when faced with the nonfocal hazard in Study 1, as a function of the numeric and verbal information they received from purported peers about the focal hazard. Error bars indicate ±1 standard errors.

In short, the analyses of behaviour toward nonfocal hazards suggest that participants did indeed interpret peers' verbal statements about one hazard as implying a comparison with the other hazard. For example, when people read that peers believed the focal hazard attacked "a lot" and "a bunch of times," they seem to take this as a sign that the nonfocal hazard did not attack very often.

Likelihood estimates about focal and nonfocal hazards

How did peers' information about the focal hazard influence participants' estimates about the same focal hazard? The overall patterns of means (see Figure 4 and Table 1 for cell statistics) revealed clear compatibility effects (Slovic, Griffin, & Tversky, 1990). That is, average numeric estimates were significantly affected by the numeric-estimate factor, $F(1, 202) = 10.69$, $p = .001$, $\eta_p^2 = .05$, but not the verbal-estimate factor, $F(2, 202) = 0.98$, $p = .38$, $\eta_p^2 = .01$. However, average verbal estimates were significantly affected by the verbal-estimate factor, $F(2, 202) = 9.46$, $p < .001$, $\eta_p^2 = .09$, but not the numeric-estimate factor, $F(1, 202) = 0.21$, $p = .65$, $\eta_p^2 = .001$.

More interesting is the question of how peers' estimates about the focal hazard influenced participants' estimates about the other hazard (see Figure 5). A 2 (numeric estimate) × 3 (verbal estimate) ANOVA on numeric estimates for the nonfocal hazard revealed a marginally significant main effect for peers' numeric estimates, $F(1, 202) = 3.90$, $p = .05$, $\eta_p^2 = 0.02$.

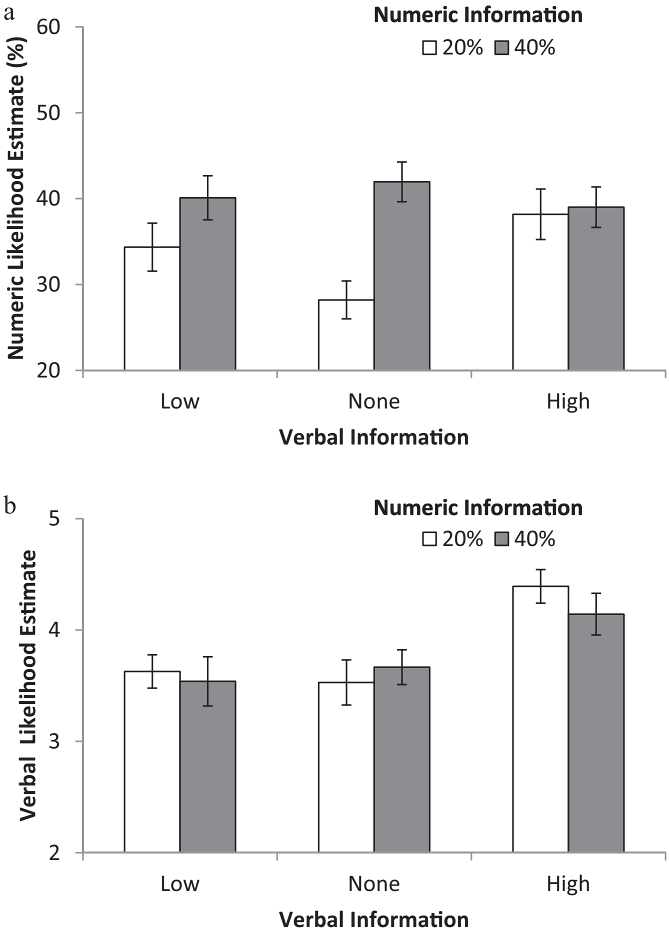

Figure 4. Participants' mean numeric (a) and verbal (b) estimates of the likelihood of a focal hazard attacking in Study 1, as a function of the numeric and verbal information they received from purported peers about the focal hazard. Error bars indicate ±1 standard errors.

Participants gave higher estimates for the nonfocal hazard when the focal hazard was said to attack 40% rather than 20% of the time (see Table 1 for full descriptive statistics). This is a collateral effect, but not an inverse one. Instead, the direction of the effect is consistent with superficial anchoring or even thoughtfully using the numeric estimate about the focal hazard as informative about the absolute likelihood of the nonfocal hazard. However, the effect of the verbal-estimate factor was significant and in an inverse direction, $F(2, 202) = 9.32$, $p < .001$, $\eta_p^2 = .08$. Participants gave higher numeric likelihood estimates regarding nonfocal hazards when the focal hazard had been described with phrases suggesting a low likelihood of attack than a high

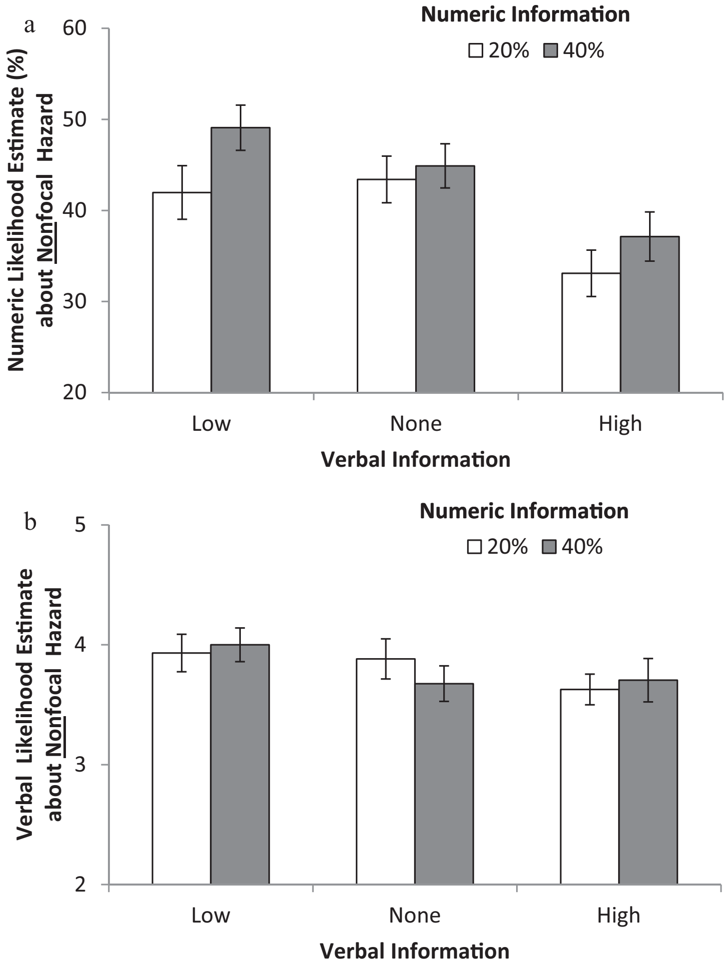

Figure 5. Participants' mean numeric (a) and verbal (b) estimates of the likelihood of a nonfocal hazard attacking in Study 1, as a function of the numeric and verbal information they received from purported peers about the focal hazard. Error bars indicate ±1 standard errors.

likelihood of attack ($p < .001$). Other Tukey comparisons revealed that the mean for the no-estimate condition, which was intermediate relative to the other two conditions, was statistically different from the high-estimate condition ($p = .002$) but not the low estimate condition ($p = .86$). The interaction was not significant, $F(2, 202) = 0.58$, $p = .56$, $\eta_p^2 = .01$.

The pattern of effects for participants' verbal estimates was generally weaker. A 2 (numeric estimate) × 3 (verbal estimate) ANOVA on verbal estimates about the nonfocal hazard produced neither significant main effects nor a significant interaction (all $ps > .14$).

The impact of individual differences

The individual difference measures were included in the study for exploratory purposes. Regression analyses revealed some moderation patterns on behaviour (e.g., lower scores on numeracy were related to larger effects of verbal expressions). However, we gain additional power for these analyses of moderators by combining data from Studies 1 and 2. Therefore, we will delay discussion of these exploratory analyses and patterns until the General Discussion.

Summary of key findings from Study 1

As expected, verbal estimates from peers had a significant influence on participants' behavioural decisions about a focal hazard, even though those participants also had numeric estimates from peers (and had experienced a learning phase that provided information about the hazards). Collateral effects on behaviour were also observed, and they were in an inverse direction. That is, when peers' verbal opinions about the attack rate of the focal hazard implied a relatively high likelihood of attack, behaviour toward the nonfocal hazard reflected more risk taking. The verbal opinions affected not only behaviour in this way; they also had an inverse effect on participants' estimates of the numeric likelihood that the nonfocal hazard would cause harm.

Study 2

In Study 1, the direct and collateral impacts of peers' verbal estimates were significant even though participants also had numeric estimates from peers. In Study 2, we tested whether these results of Study 1 would replicate under conditions where the numeric estimates were made more salient.[4] Namely, we presented participants with three estimates from peers, rather than one estimate summarised from peers. For example, in the 20% condition, participants saw estimates of 17%, 20%, and 22% ostensibly given by peers. Each estimate was in its own sentence and was attributed to a separate peer. The average of the three estimates seen by a participant was either approximately 20% or 40%, corresponding to the manipulated averages used in Study 1. Work by Obrecht, Chapman, and Gelman (2009) suggests that presenting information in this manner should increase the influence of the estimates, because there would be three encounters with data, rather than one summarised encounter.

---

[4] We do not claim that the numeric information from peers was not salient. As reported earlier, the manipulation of this information had a significant effect on participants' own numeric estimates. Also, exit questionnaire data revealed that 83.6% of the participants were able to accurately recall the peers' numeric estimate.

Method

The participants ($N = 201$) were students from introductory psychology courses at the University of Iowa. The 2 (peers' numeric estimate: 20% or 40%) × 3 (peers' verbal estimate: low, absent, or high) design was the same as for Study 1. The same blocking and counterbalancing factors were again included. Aside from the change mentioned above and some minor wording changes to the instructions, there were three other changes for Study 2. First, we modified when participants were asked the numeric and verbal likelihood questions. Participants were asked one type of likelihood question on eight occasions spread through trials 9–34 and the other type on eight occasions spread through trials 35–60 (counterbalanced). Second, we added four additional likelihood questions at the end of the test phase (verbal and numeric, about focal and nonfocal hazard). Results of these variations were largely consistent with the other likelihood judgments so are not reported further. Third, because of time constrains in early sessions for this study, the exit questionnaire and individual-difference measures were not administered for 85 participants.

Results

Preliminary analyses

There were no notable effects for gender, age, or the counterbalancing factors, so those factors are not discussed further. Rates of visiting the safe zone again dropped significantly across blocks ($p < .001$): 55.4%, 49.0%, and 44.5% for Blocks 1, 2, and 3, respectively. For reasons discussed earlier, the block factor is ignored for subsequent analyses; all key findings below are the same whether or not block is included as a factor in analyses. Exit questionnaire data revealed that 84.5% of the participants were able to recall—within 5 points—the average of the three numeric estimates attributed to peers.

Decisions about the focal hazard

We again submitted rates of visiting the safe zone on focal-hazard trials to 2 (numeric estimate) × 3 (verbal estimate) ANOVA (see Figure 6 and Table 2 for relevant means). Whereas the main effect of numeric estimates on behaviour in focal trials was marginally significant in Study 1, it was clearly not significant in this study, $F(1, 195) = 0.21$, $p = .89$, $\eta_p^2 < .001$. Contrary to Obrecht et al. (2009), the presentation of three numbers from peers rather than one average from three peers did not cause a significant increase in participants' use of the information when making decisions.

As in Study 1, the interaction was not significant ($p = .34$), but most importantly, the main effect of verbal estimates was again significant and moderately large, $F(2, 195) = 9.17$, $p < .001$, $\eta_p^2 = .09$. Participants were most likely to go to the safe zone in the high-estimate condition and least likely in the

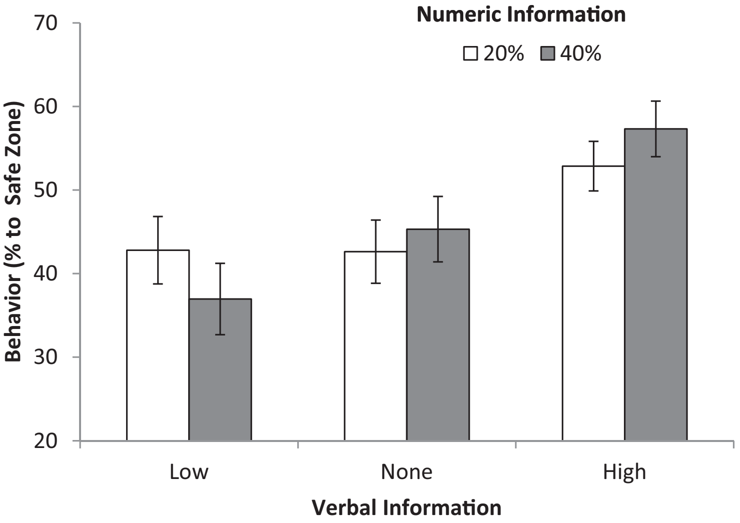

Figure 6. Participants' behavioural decisions (the mean per cent of times they went to the safe zone) when faced with the focal hazard in Study 2, as a function of the numeric and verbal information they received from purported peers about the focal hazard. Error bars indicate ±1 standard errors.

Table 2. Descriptive statistics for behaviour, numeric estimates, and verbal estimates for both focal and nonfocal hazards in Study 2.

	Verbal condition							
Dependent variable	Low		None		High		Total	
Numeric condition	M	SD	M	SD	M	SD	M	SD
Focal, behaviour								
20%	42.8	22.5	42.6	21.7	52.9	17.5	46.3	20.9
40%	37.0	24.9	45.3	22.1	57.3	20.0	46.8	23.7
Total	39.7	23.8	43.9	21.8	55.1	18.8	46.5	22.4
Focal, NLE								
20%	33.8	13.2	29.5	12.9	34.3	15.2	32.5	13.9
40%	33.3	13.7	40.4	10.7	41.2	18.9	38.3	15.2
Total	33.6	13.3	34.8	13.0	37.8	17.4	35.5	14.8
Focal, VLE								
20%	3.1	1.0	3.5	1.0	4.2	0.9	3.6	1.0
40%	3.5	0.9	3.8	1.0	4.2	1.3	3.8	1.1
Total	3.3	0.9	3.7	1.0	4.2	1.1	3.7	1.1
Nonfocal, behaviour								
20%	62.7	25.9	55.4	18.6	45.0	21.1	54.0	22.9
40%	59.3	23.3	47.2	21.9	47.6	24.0	51.4	23.6
Total	60.9	24.4	51.3	20.5	46.3	22.5	52.7	23.2
Nonfocal, NLE								
20%	42.4	16.8	37.3	18.1	29.4	15.8	36.1	17.6
40%	43.4	17.3	45.6	14.6	37.5	19.3	42.0	17.5
Total	42.9	16.9	41.4	16.9	33.5	18.0	39.1	17.7
Nonfocal, VLE								
20%	4.2	1.0	3.7	0.8	3.7	1.1	3.9	1.0
40%	4.1	1.1	4.1	0.8	3.9	1.1	4.0	1.0
Total	4.2	1.1	3.9	0.9	3.8	1.1	4.0	1.0

Note: Behaviour = % of times going to safe zone when hazard was present. NLE = Numeric Likelihood Estimate (0%–100%). VLE = Verbal Likelihood Estimate (score from 1 to 7). Condition information refers to what participants were told that peers had said about the focal hazard (never the nonfocal hazard).

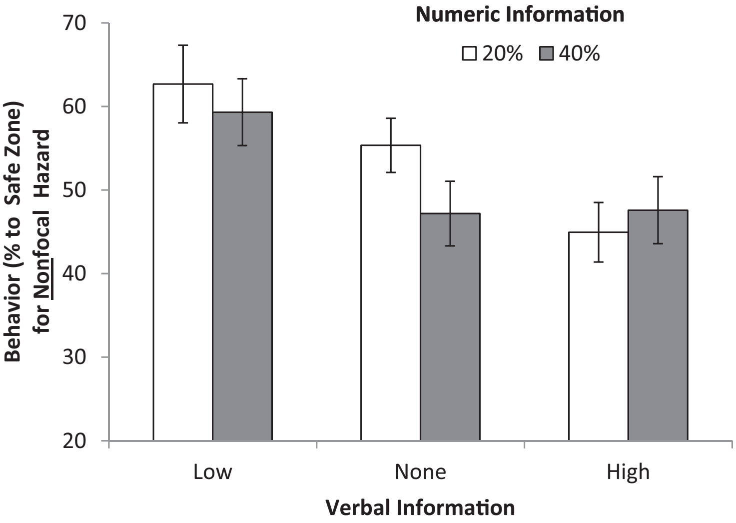

Figure 7. Participants' behavioural decisions (the mean per cent of times they went to the safe zone) when faced with the nonfocal hazard in Study 2, as a function of the numeric and verbal information they received from purported peers about the focal hazard. Error bars indicate ±1 standard errors.

low-estimate condition ($p < .001$ for the difference). Other Tukey comparisons revealed that safe-zone use in the no-estimate condition was significantly lower than in the high-estimate condition ($p = .008$) but not different from in the low estimate condition ($p = .56$).

Decisions about the nonfocal hazard

Not only was the direct effect of verbal estimates replicated, but the collateral effects were as well. In a 2 × 3 ANOVA on safety-zone rates in nonfocal trials, the numeric-estimate main effect and interaction were not significant (both ps $> .35$) but the main effect of the verbal estimate factor was significant, $F(2, 195) = 7.40$, $p < .001$, $\eta_p^2 = 0.07$. The direction of the effect on the nonfocal hazard was again the inverse of that for the focal hazard (see Figure 7 and Table 2). Participants were more likely to go to the safe zone for the nonfocal hazard when the focal hazard had been described with phrases suggesting a low rather than high likelihood of attack ($p < .001$). Other Tukey comparisons revealed that the mean for the no-estimate condition, which was intermediate relative to the other two conditions, was statistically different from the low-estimate condition ($p < .05$) but not the high estimate condition ($p = .40$).

Likelihood estimates about the focal and nonfocal hazards

For likelihood estimates about focal hazards, the same compatibility effects were observed as in Study 1 (see Figure 8 and Table 2). Participants' numeric

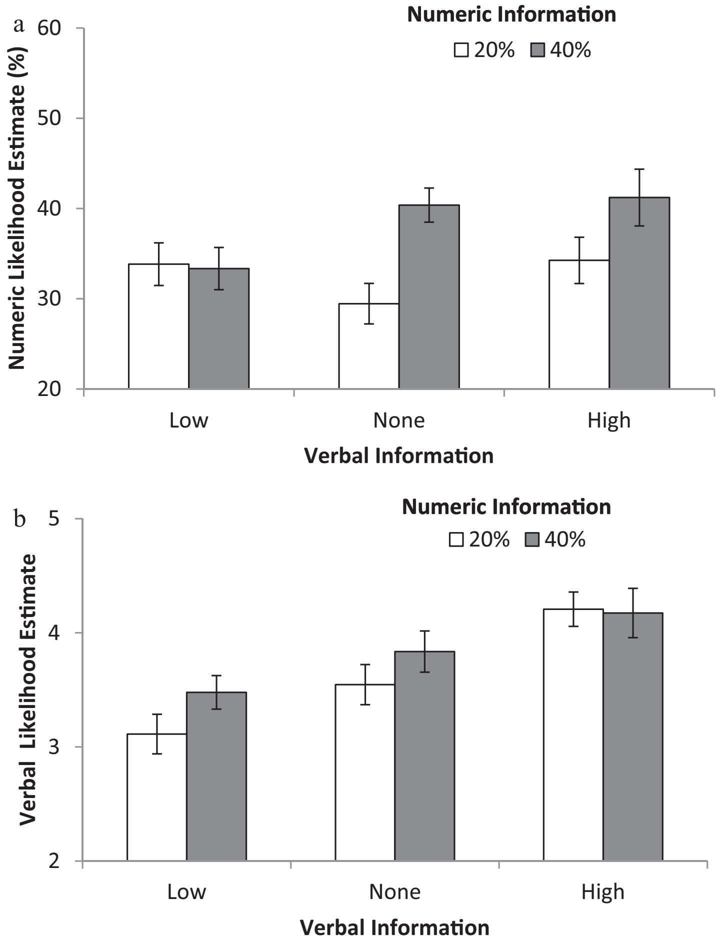

Figure 8. Participants' mean numeric (a) and verbal (b) estimates of the likelihood of a focal hazard attacking in Study 2, as a function of the numeric and verbal information they received from purported peers about the focal hazard. Error bars indicate ±1 standard errors.

estimates were significantly affected by the numeric-estimate factor, $F(1, 195) = 8.10$, $p = .005$, $\eta_p^2 = .04$, but not by the verbal-estimate factor, $F(2, 195) = 1.48$, $p = .23$, $\eta_p^2 = .02$. Average verbal estimates were significantly affected by the verbal-estimate factor, $F(2, 195) = 13.15$, $p < .001$, $\eta_p^2 = .12$, but not the numeric-estimate factor, $F(1, 195) = 2.01$, $p = .15$, $\eta_p^2 = .01$.

More important, the pattern of collateral effects on numeric estimates about nonfocal hazards was the same as found in Study 1. There was a significant main effect for peers' numeric estimates, again in a direction consistent with anchoring rather than in an inverse direction, $F(1, 195) = 5.72$, $p = .02$, $\eta_p^2 = 0.03$. Meanwhile, there was again an inverse influence on numeric

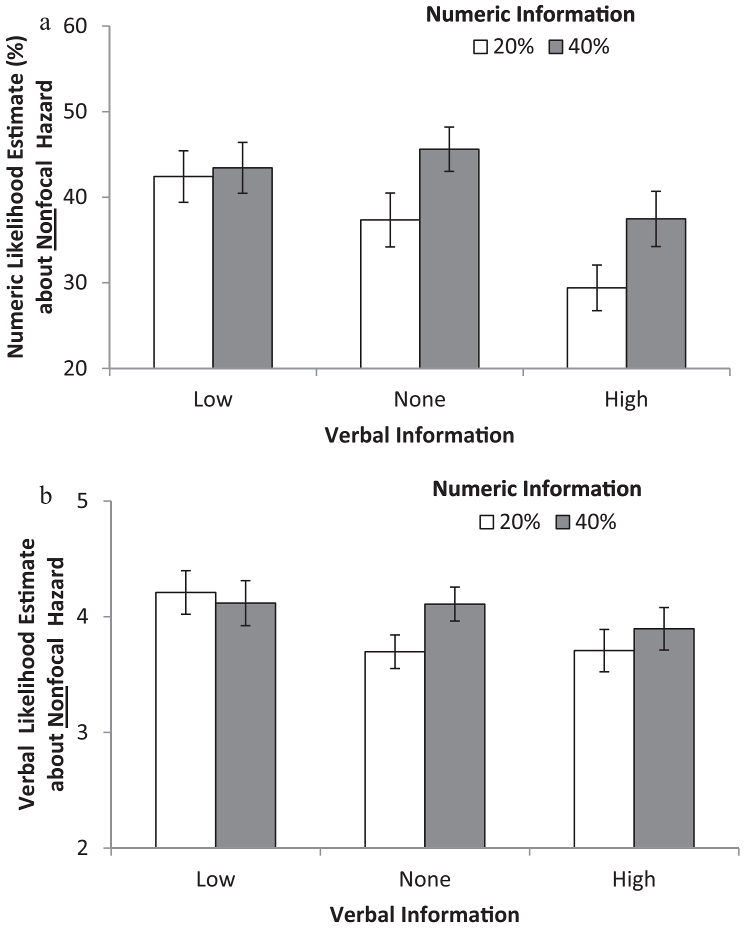

Figure 9. Participants' mean numeric (a) and verbal (b) estimates of the likelihood of a nonfocal hazard attacking in Study 2, as a function of the numeric and verbal information they received from purported peers about the focal hazard. Error bars indicate ±1 standard errors.

estimates that was triggered by peers' verbal estimates, $F(2, 195) = 6.14$, $p = .003$, $\eta_p^2 = .06$. The interaction was not significant ($p = .39$; see Figure 9 and Table 2 for means).

The pattern of effects for participants' verbal estimates about the nonfocal hazard was again generally weaker. A 2 × 3 ANOVA produced neither significant main effects nor a significant interaction (all $ps > .10$).

The impact of individual differences

Again, see the later part of the General Discussion, where we present our exploratory analyses of individual differences based on a combined data-set (Studies 1 and 2).

Summary for findings from study 2

Study 2 replicated the key findings from Study 1. Peers’ verbal estimates had a significant influence on participants’ behavioural decisions even though participants also had three numeric estimates from peers. The collateral effects of peers’ verbal estimates were again observed and were again inverse to the effects found regarding focal hazards. As in Study 1, this inverse, collateral effect was observed not only on behavioural responses but also on participants’ numeric likelihood estimates about the nonfocal hazard. The only notable difference between Studies 1 and 2 was that, perhaps surprisingly, the impact of peers’ numeric estimates on behaviour was not significant in Study 2 (it was marginal in Study 1).

General discussion

In two studies, we used a novel task—the DART—to assess the influence of verbal likelihood information in dual-mode communications (i.e., when a speaker provides information using both verbal and numeric expressions). DART allowed us to test, with repeated trials, the impact of this information on behavioural decisions. This included behaviour toward the hazard about which the peer's likelihood opinion was directed (the focal hazard) as well as another hazard in the same context (the nonfocal hazard).

Decisions about the focal hazard

As documented in a rich literature, verbal expressions of probability often have vague intended meanings and are open to substantial variability in interpretations across persons, cultures, and contexts (e.g., Beyth-Marom, 1982; Brun & Teigen, 1988; Budescu & Wallsten, 1995; Harris et al., 2013; Wallsten et al., 1986). Furthermore, people tend to prefer to receive probabilistic information from others in numerical rather than verbal form (Erev & Cohen, 1990; Olson & Budescu, 1997; Shaw & Dear, 1990; see also Gurmankin et al., 2004a, 2004b). Given these facts, one might expect that verbal likelihood expressions—when uttered by peers or experts—would be rather limited in their main-effect influences on a people's decisions and behaviours, particularly when people are given precise numeric information from the same sources. Our prediction, however, was that peers’ verbal estimates would have a significant impact on decisions about a focal hazard even when people were also given numeric estimates. The findings from both studies were consistent with our prediction. Participants were more likely to go to the safe zone when peers’ verbal terms describing the attack rate suggested a high rather than low likelihood of attack, even though participants had numeric information from peers. Apparently, people's desire for the precision of a numeric estimate does not mean they will exclusively rely on it when available. Instead, they continue to be influenced by the connotations or denotations of the verbal estimates. For example, hearing an expert say that a medicine produces side effects “only rarely, or about 20% of the time” is likely to have a different impact on relevant decisions and behaviour than hearing “occasionally, or about 20% of the time”.

This may happen because numerical statistics, although precise, do not inherently convey an intuitive meaning nor always have compelling impact on decisions (Borgida & Nisbett, 1977; Hsee, 1996; Windschitl et al., 2002; Zikmund-Fisher, 2013; Zikmund-Fisher et al., 2004). To the extent that they do not have compelling impact, then the probabilities implied by the low vs. high verbal expressions may have their own independent influence on decision-making. More generally, one could view a numeric likelihood estimate, such as the peers’ estimates in our studies, as an imperfect signal for objective likelihood. In this case, a second estimate—whether it is verbal or numeric or in some other form—could be helpful for best approximating objective likelihood, assuming the error associated with one estimate is independent from the error associated with the other estimate. Consequently, our experiments do not allow for conclusions about whether an independent influence from verbal estimates should necessarily be considered irrational or generally problematic for achieving optimal performance. Nonetheless, it is important to recall that peers’ verbal expressions did not significantly affect participants’ numeric estimates of likelihood in either experiment. These null patterns suggest that the impact of peers’ verbal expressions on participant behaviour is not explained by a simple process in which participants were using verbal estimates as a second signal for establishing their best estimate of objective, numeric likelihood. Other dynamics were likely at play.

To address these possible dynamics, we return to the issue of how people interpret and draw meaning from numeric likelihood information. Research suggests that interpretations of even a very precise numeric estimate are malleable, and there may be various ways in which people use contextual information to draw meaning or an interpretation from a specific number (e.g., Bigman, 2014; Fagerlin et al., 2007; Flugstad & Windschitl, 2003; Klein, 1997; McKenzie & Nelson, 2003; Windschitl & Weber, 1999; Windschitl et al., 2002). In the present studies, the verbal estimates from peers might have presented a rather strong contextual cue for interpreting the numeric information. Verbal expressions of likelihood tend to contain directional signals from a speaker (e.g., Sanford & Moxey, 2003; Teigen & Brun, 1995), and the verbal expressions we presented to participants in our studies certainly contained directional information about how peers felt about the danger of attack by the given hazard. The directional information appears to have influenced how participants thought about and behaved in response to the uncertainty associated with a focal hazard, without significantly shifting numeric estimations of the likelihood of a hazard attacking.

In sum, one key conclusion from this work is that the impact of numerical expressions of uncertainty do not overwhelm or heavily constrain the impact of verbal expressions of uncertainty. This is a conclusion based on work that had decisions as the main dependent measure, but it also comports with related research measuring people's numeric interpretations of verbal likelihood expressions. Budescu et al. (2009) examined how numeric interpretations of verbal probability terms in a document from the Intergovernmental Panel on Climate Change were affected when those terms were accompanied by numeric guidelines (see also Budescu, Por, & Broomell, 2012). Although they found that the numbers led to a reduction in the inter-individual variability of interpretations, substantial variability persisted even when the guidelines appeared immediately adjacent to the verbal probability terms. These findings and ours highlight an important caveat to the recommendation that important communications that contain verbal expressions of uncertainty about risks should also contain numeric estimates for clarity (e.g., Budescu et al., 2009; Harris et al., 2013; Juanchich, Sirota, & Butler, 2012; Visschers, Meertens, Passchier, & de Vries, 2009; see also Gurmankin, Baron, & Armstrong, 2004b). Namely, this recommendation should not be confused with the suggestion that the numeric information will remove all or even most bias and problematic variability.

Decisions about a nonfocal hazard: collateral inverse effects of verbal expressions

Verbal expressions not only had an impact on participants' decisions about the focal hazard, they also had a collateral impact. This impact was in an inverse direction. In both studies, participants were more likely to go to the safe zone for the nonfocal hazard when the focal hazard had been described by peers with phrases suggesting a low rather than high likelihood of attack. A similar pattern was observed regarding participants' numeric likelihood estimates. Participants gave higher likelihood estimates regarding an attack from a nonfocal hazard when the focal hazard had been described with phrases suggesting a low rather than high likelihood of attack.

Prior research has examined how people not only select but also decode subjective evaluations (Higgins & Lurie, 1983; Higgins & Stangor, 1988; Kobrynowicz & Biernat, 1997). Consider a case in which a speaker mentions to a listener that a particular bicycle racer is “fantastic” in 10-mile time trials. In decoding what “fantastic” means, the listener will consider the standards that the speaker is likely to have used when deciding to use that language. If the listener knows that the speaker is typically watching world-class cyclists, the standard of comparison will be assumed to be very high. If the listener knows that the speaker is typically watching amateurs, the standard of comparison will be assumed to be much lower. Critically, if asked to estimate a numeric time for this “fantastic” bike racer, the numeric estimate of the listener would be a faster time in the former case than the latter case. The results of our studies suggest that participants likely engaged in a similar but slightly different line of reasoning. They appear to have assumed that the nonfocal hazard was a standard that peers had in mind when evaluating the attack rate or deciding how to characterise the attack rate of the focal hazard. Therefore, if the peers characterised the focal hazard as attacking “often,” which seems like a comparative term pointing in the high-attack direction, then the other hazard must have attacked less often.

There are innumerable everyday contexts in which verbal expressions aimed at one hazard could affect interpretation of another hazard. For example, imagine a zoo intern who must relocate two potentially dangerous animals. Hearing a fellow intern say “I doubt the grey one will try to bite you” may not only reduce the intern's caution when handling the grey animal, but it may also increase concern when handling the tan animal. The collateral, inverse effect of verbal likelihood phrases might raise an important lesson for other contexts. For example, if public health officials use evaluative terminology to characterise the vulnerability of a particular group as high, this might reduce the perceived vulnerability of people in a comparable but different group (for related work see Bigman, 2014; Windschitl et al., 2002).

Again, the manipulation of peers’ verbal expressions had no significant impact directly on people's numeric estimates of the focal hazards’ likelihood of attacking. However, verbal expressions did have a significant effect on people's estimates of the other hazards’ likelihood of attacking (the collateral, inverse effects). It may seem puzzling as to why verbal information lacks a direct impact but enjoys a collateral impact on the same dependent variable. Perhaps this is because participants were always given a peer's numeric estimate about the focal hazard but not about the nonfocal hazard. The peers’ numeric estimate may have anchored the participants’ own numeric estimate about the focal hazard, but not the nonfocal hazard (or at least not to the same extent). Without a strong anchor influencing participants’ responses to the nonfocal hazard, decoding of what the verbal expression suggests about the nonfocal hazard might be quite influential (for more on decoding in other contexts, see Higgins & Lurie, 1983; Higgins & Stangor, 1988; Kobrynowicz & Biernat, 1997).

Limitations

Our studies have a few notable limitations relevant to the collateral and/or direct effects. First, none of our conditions omitted numeric estimates entirely, so we did not test whether verbal expressions would have the inverse, collateral impact on behaviour when not accompanied by numeric estimates. There is no theoretical reason to assume that the inverse effect on behaviour would fail to work in the absence of numeric estimates, but our studies offer no empirical verification of this assumption. Second, we did not test how the collateral effect might operate when there are three or more hazards. We can only presume that the inverse effect on behaviour toward a nonfocal hazard is greater when it is the only nonfocal hazard within the particular context than when there are many nonfocal hazards. Additionally, we could only speculate about whether the inverse impact drops off linearly or suddenly as the number of nonfocal hazards increase from one. Third, we did not empirically isolate the key ingredient that is causing a verbal expression about one hazard to have an inverse impact on beliefs about another. We have suggested that the implied directionality and comparative nature of an evaluative term, such as "often," is crucial, but this does not mean that all verbal terms would produce inverse collateral effects. Furthermore, there remains the possibility that statements containing numeric estimates sometimes may suggest a directionality and comparative inference that triggers an inverse effect on behaviour toward a nonfocal hazard. It is also possible that part of the influence that the verbal likelihood phrases had in our studies is due to inferences participants made about peer's beliefs about how to behave toward hazards. Finally, in our choice of verbal terms to use in our studies, directionality is confounded with level of implied probability because the words that suggested high probability are also pointed toward occurrences rather than non-occurrences (Teigen & Brun, 2003b). This is a natural confound but one that could be experimentally disentangled to more confidently isolate directionality as the key causal factor for our results.

The impact of individual differences

Participants in both studies completed a numeracy scale (Lipkus et al., 2001), a subjective numeracy scale (Fagerlin et al., 2007), Need for Cognition Scale (Cacioppo et al., 1984), and a dispositional optimism scale (Scheier et al., 1994). We combined the data from Studies 1 and 2 to maximise the power of our exploratory analyses of the impact of these individual differences on sensitivity to peers' estimates. In these analyses, we initially coded for study, but this rarely affected results. We found no noteworthy effects of dispositional optimism on any dependent variable or of the other three individual-difference variables on participants' likelihood estimates. What did emerge was evidence that numeracy and need for cognition moderated how verbal information from peers (not numeric information) influenced participants' behaviour regarding focal and nonfocal hazards.

For the key analyses that revealed these effects, the low-verbal-expression condition was coded as −1, the no-expression condition was 0, and the high-

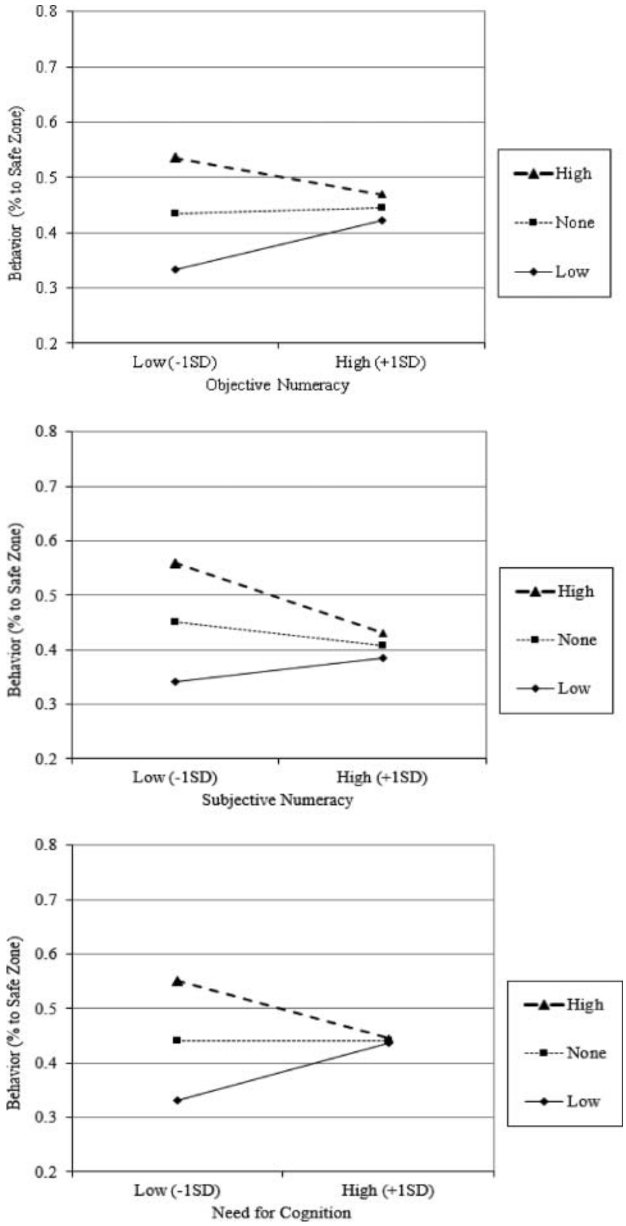

Figure 10. Representation of participants' behaviour (from regression) in *focal* trials as a function of the verbal information they received and objective numeracy (a), subjective numeracy (b), or need for cognition (c).

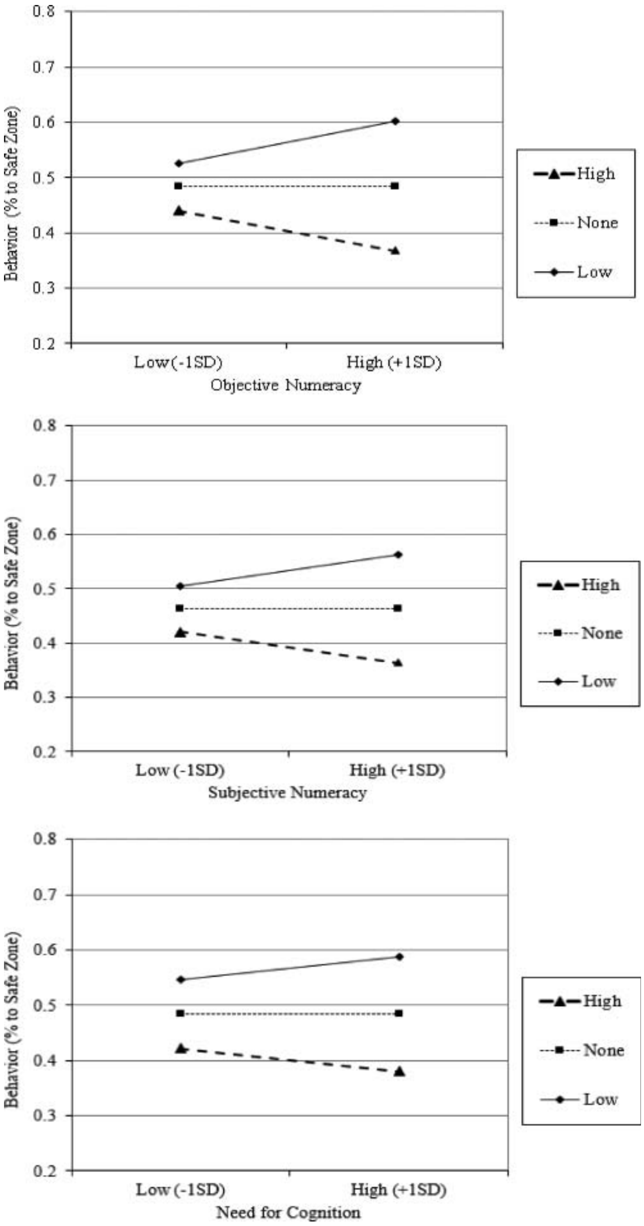

Figure 11. Representation of participants' behaviour (from regression) in *nonfocal* trials as a function of the verbal information they received and objective numeracy (a), subjective numeracy (b), or need for cognition (c).

verbal-expression condition was +1. We conducted a series of regression analyses in which this verbal-expression factor, a given individual-difference variable (mean-centred), and their respective interaction were predictors, and either behaviour during focal or nonfocal trials was the outcome variable. In separate analyses examining behaviour toward focal hazards as the outcome variable, each of the following emerged as moderators, as evidenced by significant interactions with the verbal-expression factor: objective numeracy, $t(316) = -2.76$, $p = .006$, $\beta = -0.15$, subjective numeracy, $t(267) = -2.96$, $p = .003$, $\beta = -0.17$, and need for cognition, $t(316) = -3.75$, $p < .001$, $\beta = -0.20$.[5] In all three cases, lower scores on the individual-difference variable were associated with a greater impact of verbal expressions on behaviour in focal trails (see Figure 10(a)–10(c)). Thus, the effect of verbal phrases on behaviour in focal trials was greater for participants who were low on need for cognition and for those who were low in numeracy (both for subjective and objective measures).

Meanwhile, for behaviour on the non-focal trials, the objective numeracy skill produced a significant interaction with the verbal-expression factor, $t(316) = -2.82$, $p = .005$, $\beta = -0.15$. It was the highly numerate participants who were more affected by the verbal information (see Figure 11(a)–11(c)). The interaction patterns involving subjective numeracy, $t(267) = -1.93$, $p = .06$, $\beta = -0.11$, was marginally significant and the interaction with need for cognition, $t(316) = -1.45$, $p = .149$, $\beta = -0.08$, was not significant.

These findings were based on exploratory analyses and call for additional validation, but they do suggest a pair of speculations relevant to the literature on numeracy (e.g., Betsch, Haase, Renkewitz, & Schmid, 2015; Peters et al., 2006; Peters, Hart, Tusler, & Fraenkel, 2014). First, they suggest that low numerate individuals may be especially influenced in an obvious direction by verbal information. That is, a peer's expression about a threat will have a direct effect on behaviour toward that threat. Second, high numerate individuals may be especially influenced in a nonobvious direction by verbal information. When they hear that a peer described a threat as dangerous, they are more likely than low-numerate individuals to draw inferences about what this means regarding the dangerousness of another threat (in an inverse direction).[6]

Conclusion

We used a new task, the DART, to study how decisions within a risky environment would be influenced by likelihood information from peers. We found

---

[5] The correlation between scores for objective and subjective numeracy was $r = .39$ ($p < .001$), between objective numeracy and need for cognition was $r = .16$ ($p < .01$), and between subjective numeracy and need for cognition was $r = .37$ ($p <$ 001).

[6] A challenge for this speculation is that numeracy did not significantly interact with the verbal estimate factor to influence participants' numeric or verbal estimates.

that verbal expressions of likelihood, although known for being imprecise and open to systematic and random variability in interpretation, were clearly influential even in the presence of numeric estimates. That is, even when people were given peers' numeric likelihood estimates about a hazard, the same peers' verbal estimates influenced participants' behaviour. Their influence did not end there, however. Verbal estimates also had collateral, inverse influences on behaviour and estimates about other hazards in the same context. Both the direct and collateral effects of verbal estimates that were detected here should serve as a notice for speakers (and writers) attempting to communicate a level of certainty about an event. Some speakers might believe that their numeric estimate of likelihood provides a precise bottom line that supersedes the verbal expressions that they also offer, but this would be mistaken. Verbal expressions that are also uttered have the potential for additional and unanticipated influences on listeners

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the National Science Foundation [grant number SES 09-61252].

References

Beyth-Marom, R. (1982). How probable is probable? A numerical translation of verbal probability expressions. *Journal of Forecasting, 1*, 257–269. doi:10.1002/for.3980010305

Betsch, C., Haase, N., Renkewitz, F, & Schmid, P. (2015). The narrative bias revisited: What drives the biasing influence of narrative information on risk perceptions? *Judgment and Decision Making, 10*, 241–264.

Biernat, M., Manis, M., & Nelson, T. E. (1991). Stereotypes and standards of judgment. *Journal of Personality and Social Psychology, 60*, 485–499. doi: 10.1037/0022-3514.60.4.485

Bigman, C. A. (2014). Social comparison framing in health news and its effect on perceptions of group risk. *Health Communication, 29*, 267–280. doi:10.1080/10410236.2012.745043

Bilgin, B., & Brenner, L. (2013). Context affects the interpretation of low but not high numerical probabilities: A hypothesis testing account of subjective probability. *Organizational Behavior and Human Decision Processes, 121*, 118–128. doi: 10.1016/j.obhdp.2013.01.004

Borgida, E., & Nisbett, R. E. (1977). The differential impact of abstract vs. concrete information on decisions. *Journal of Applied Social Psychology, 7*, 258–271. doi: 10.1111/j.1559-1816.1977.tb00750.x

Brun, W., & Teigen, K. H. (1988). Verbal probabilities: Ambiguous, context-dependent, or both? *Organizational Behavior and Human Decision Processes, 41*, 390–404. doi:10.1016/0749-5978(88)90036-2

Budescu, D. V., Broomell, S., & Por, H. H. (2009). Improving communication of uncertainty in the reports of the intergovernmental panel on climate change. *Psychological Science, 20*, 299–308. doi: 10.1111/j.1467-9280.2009.02284.x

Budescu, D. V., Por, H. H., & Broomell, S. B. (2012). Effective communication of uncertainty in the IPCC reports. *Climatic Change, 113*, 181–200. doi:10.1007/s10584-011-0330-3

Budescu, D. V., Por, H.-H., Broomell, S. B., & Smithson, M. (2014). The interpretation of IPCC probabilistic statements around the world. *Nature Climate Change, 4*, 508–512. doi:10.1038/nclimate2194

Budescu, D. V., & Wallsten, T. S. (1995). Processing linguistic probabilities: General principles and empirical evidence. *Psychology of Learning and Motivation, 32*, 275–318. doi:10.1016/S0079-7421(08)60313-8

Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment, 48*, 306–307. doi:10.1207/s15327752jpa4803—13

Erev, I., & Cohen, B. L. (1990). Verbal versus numerical probabilities: Efficiency, biases, and the preference paradox. *Organizational Behavior and Human Decision Processes, 45*, 1–18. doi:10.1016/0749-5978(90)90002-Q

Fagerlin, A., Zikmund-Fisher, B. J., & Ubel, P. A. (2007). "If I'm better than average, then I'm ok?": Comparative information influences beliefs about risk and benefits. *Patient Education and Counseling, 69*, 140–144, doi:10.1016/j.pec.2007.08.008

Fagerlin, A., Zikmund-Fisher, B. J., Ubel, P. A., Jankovic, A., Derry, H. A., & Smith, D. M. (2007). Measuring numeracy without a math test: Development of the Subjective Numeracy Scale. *Medical Decision Making, 27*, 672–680. doi:10.1177/0272989×07304449

Fischer, K., & Jungermann, H. (1996). Rarely occurring headaches and rarely occurring blindness: Is rarely = rarely? *Journal of Behavioral Decision Making, 9*, 153–172. doi:10.1002/(SICI)1099-0771(199609)9:3<153::AID-BDM222>3.0.CO;2-W

Flugstad, A. R., & Windschitl, P. D. (2003). The influence of reasons on interpretations of probability forecasts. *Journal of Behavioral Decision Making, 16*, 107–126. doi:10.1002/bdm.437

Gigerenzer, G., Hertwig, R., Van Den Broek, E., Fasolo, B., & Katsikopoulos, K. V. (2005). "A 30% chance of rain tomorrow": How does the public understand probabilistic weather forecasts? *Risk Analysis, 25*, 623–629. doi:10.1111/j.1539-6924.2005.00608.x

Gurmankin, A. D., Baron, J., & Armstrong, K. (2004a). Intended message versus message received in hypothetical physician risk communications: Exploring the gap. *Risk Analysis, 24*, 1337–1347. doi:10.1111/j.0272-4332.2004.00530.x

Gurmankin, A. D., Baron, J., & Armstrong, K. (2004b). The effect of numerical statements of risk on trust and comfort with hypothetical physician risk communication. *Medical Decision Making, 24*, 265–271. doi:10.1177/0272989X04265482

Harris, A. J. L., & Corner, A. (2011). Communicating environmental risks: Clarifying the severity effect in interpretations of verbal probability expressions. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 37*, 1571–1578. doi:10.1037/a0024195

Harris, A. J. L., Corner, A., Xu, J., & Du, X. (2013). Lost in translation? Interpretations of the probability phrases used by the Intergovernmental Panel on Climate Change in China and the UK. *Climatic Change, 121*, 415–425. doi:10.1007/s10584-013-0975-1

Higgins, E. T., & Lurie, L. (1983). Context, categorization, and recall: The “change-of-standard” effect. *Cognitive Psychology, 15*, 525–547. doi:10.1016/0010-0285(83)90018-X

Higgins, E. T., & Stangor, C. (1988). A “change-of-standard” perspective on the relations among context, judgment, and memory. *Journal of Personality and Social Psychology, 54*, 181–192. doi:10.1037/0022-3514.54.2.181

Honda, H., & Yamagishi, K. (2006). Directional verbal probabilities: Inconsistencies between preferential judgments and numerical meanings. *Experimental Psychology, 53*, 161–170. doi:10.1027/1618-3169.53.3.161

Hsee, C. K. (1996). The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational Behavior and Human Decision Processes, 67*, 247–257. doi:10.1006/obhd.1996.0077

Hsee, C. K., & Zhang, J. (2010). General evaluability theory. *Perspectives on Psychological Science, 5*, 343–355. doi:10.1177/1745691610374586

Huttenlocher, J., & Higgins, E. T. (1971). Adjectives, comparatives, and syllogisms. *Psychological Review, 78*, 487–504. doi:10.1037/h0031858

Juanchich, M., Sirota, M., & Butler, C. L. (2012). The perceived functions of linguistic risk quantifiers and their effect on risk, negativity perception and decision making. *Organizational Behavior and Human Decision Processes, 118*, 72–81. doi:10.1016/j.obhdp.2012.01.002

Juanchich, M., Sirota, M., Karelitz, T. M., & Villejoubert, G. (2013). Can membership-functions capture the directionality of verbal probabilities? *Thinking & Reasoning, 19*, 231–247. doi:10.1080/13546783.2013.772538

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review, 93*, 136–15. doi:10.1037/0033-295X.93.2.136

Klein, W. M. (1997). Objective standards are not enough: Affective, self-evaluative, and behavioral responses to social comparison information. *Journal of Personality and Social Psychology, 72*, 763–774. doi:10.1037/0022-3514.72.4.763

Kobrynowicz, D., & Biernat, M. (1997). Decoding subjective evaluations: How stereotypes provide shifting standards. *Journal of Experimental Social Psychology, 33*, 579–601. doi:10.1006/jesp.1997.1338

Levin, I. P., Schneider, S. L., & Gaeth, G. J. (1998). All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes, 76*, 149–188. doi:10.1006/obhd.1998.2804

Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making, 21*, 37–44. doi:10.1177/0272989X0102100105

McKenzie, C. R. M., & Nelson, J. D. (2003). What a speaker's choice of frame reveals: Reference points, frame selection, and framing effects. *Psychonomic Bulletin & Review, 10*, 596–602. doi:10.3758/BF03196520

Moxey, L. M., & Sanford, A. J. (2000). Communicating quantities: A review of psycholinguistic evidence of how expressions determine perspectives. *Applied Cognitive Psychology, 14*, 237–255. doi:10.1002/(SICI)1099-0720(200005/06)14:3<237::AID-ACP641>3.0.CO;2-R

Obrecht, N. A., Chapman, G. B., & Gelman, R. (2009). An encounter frequency account of how experience affects likelihood estimation. *Memory Cognition, 37*, 632–643. doi:10.3758/MC.37.5.632

Olson, M. J., & Budescu, D. V. (1997). Patterns of preference for numerical and verbal probabilities. *Journal of Behavioral Decision Making, 10*, 117–131. doi:10.1002/(SICI)1099-0771(199706)10:2<117::AID-BDM251>3.0.CO;2-7

Parducci, A. (1965). Category judgment: A range-frequency model. *Psychological Review, 72*, 407–418. doi:10.1037/h0022602

Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological Science, 17*, 407–413. doi:10.1111/j.1467-9280.2006.01720.x

Peters, E., Hart, P. S., Tusler, M., & Fraenkel, L. (2014). Numbers matter to informed patient choices: A randomized design across age and numeracy levels. *Medical Decision Making, 34*, 430–442. doi:10.1177/0272989×13511705

Piercey, M. D. (2009). Motivated reasoning and verbal vs. numerical probability assessment: Evidence from an accounting context. *Organizational Behavior and Human Decision Processes, 108*, 330–341. doi:10.1016/j.obhdp.2008.05.004

Sanford, A. J., & Moxey, L. M. (2003). New perspectives on the expression of quantity. *Current Directions in Psychological Science, 12*, 240–243. doi:10.1046/j.0963-7214.2003.01270.x

Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A reevaluation of the life orientation test. *Journal of Personality and Social Psychology, 67*, 1063–1078. doi:10.1037/0022-3514.67.6.1063

Shaw, N. J., & Dear, P. R. (1990). How do parents of babies interpret qualitative expressions of probability? *Archives of Disease in Childhood, 65*, 520–523. doi:10.1136/adc.65.12.1379

Sher, S., & McKenzie, C. R. M. (2014). Options as information: Rational reversals of evaluation and preference. *Journal of Experimental Psychology: General, 143*, 1127–1143. doi:10.1037/a0035128

Slovic, P., Griffin, D., & Tversky, A. (1990). Compatibility effects in judgment and choice. In R. M. Hogarth (Ed.), *Insights in decision making: A tribute to Hillel J. Einhorn* (pp. 5–27). Chicago, IL: The University of Chicago Press.

Teigen, K. H., & Brun, W. (1995). Yes, but it is uncertain: Direction and communicative intention of verbal probabilistic terms. *Acta Psychologica, 88*, 233–258. doi:10.1016/0001-6918(93)E0071-9

Teigen, K. H., & Brun, W. (2000). Ambiguous probabilities: When does p = 0.3 reflect a possibility, and when does it express a doubt? *Journal of Behavioral Decision Making, 13*, 345–362. doi:10.1002/1099-0771(200007/09)13:3<345::AID-BDM358>3.0.CO;2-U

Teigen, K. H., & Brun, W. (2003a). Verbal probabilities: A question of frame? *Journal of Behavioral Decision Making, 16*, 53–72. doi:10.1002/bdm.432

Teigen, K. H., & Brun, W. (2003b). Verbal expressions of uncertainty and probability. In D. Hardman & L. Macchi (Eds.), *Thinking: Psychological perspectives on reasoning, judgment and decision making* (pp. 125–145). Chichester: John Wiley & Sons. doi:10.1002/047001332X.ch7

Visschers, V. H. M., Meertens, R. M., Passchier, W. W. F., & de Vries, N. N. K. (2009). Probability information in risk communication: A review of the research literature. *Risk Analysis, 29*, 267–287. doi:10.1111/j.1539-6924.2008.01137.x

Vlaev, I., Chater, N., Stewart, N., & Brown, G. D. A. (2011). Does the brain calculate value? *Trends in Cognitive Sciences, 15*, 546–554. doi:10.1016/j.tics.2011.09.008

Wallsten, T. S., Budescu, D. V., Rapoport, A., Zwick, R., & Forsyth, B. (1986). Measuring the vague meanings of probability terms. *Journal of Experimental Psychology: General, 115*, 348–365. doi:10.1037/0096-3445.115.4.348

Weber, E. U. (1994). From subjective probabilities to decision weights: The effect of asymmetric loss functions on the evaluation of uncertain outcomes and events. *Psychological Bulletin, 115*, 228–242. doi:10.1037/0033-2909.115.2.228

Weber, E. U., & Hilton, D. J. (1990). Contextual effects in the interpretations of probability words: Perceived base rate and severity of events. *Journal of Experimental Psychology Human Perception and Performance, 16*, 781–789. doi:10.1037/0096-1523.16.4.781

Windschitl, P. D., Martin, R., & Flugstad, A. R. (2002). Context and the interpretation of likelihood information: The role of intergroup comparisons on perceived vulnerability. *Journal of Personality and Social Psychology, 82*, 742–755. doi:10.1037/0022-3514.82.5.742

Windschitl, P. D., & Weber, E. U. (1999). The interpretation of "likely" depends on the context, but "70%" is 70%–right? The influence of associative processes on perceived certainty. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 25*, 1514–1533. doi:10.1037//0278-7393.25.6.1514

Zikmund-Fisher, B. J. (2013). The right tool is what they need, not what we have: A taxonomy of appropriate levels of precision in patient risk communication. *Medical Care Research and Review, 70*(Suppl. 1), 37S–49S. doi:10.1177/1077558712458541

Zikmund-Fisher, B. J., Fagerlin, A., & Ubel, P. A. (2004). "Is 28% good or bad?" Evaluability and preference reversals in health care decisions. *Medical Decision Making, 24*, 142–148. doi:10.1177/0272989×04263154