Using risk model judgements to better understand perceptions of synergistic risks

Ian G.J. Dawson[1], Johnnie E.V. Johnson[1], and Michelle A. Luke[2]

[1] Centre for Risk Research, University of Southampton, Southampton, SO17 1BJ, UK.

[2] School of Business, University of Sussex, Brighton, BN1 9SL, UK.

*Requests for reprints should be addressed to Ian G.J. Dawson, Centre for Risk Research, School of Management, University of Southampton, SO17 1BJ, UK. (e-mail: I.G.Dawson@Soton.ac.uk).

Abstract

Numerous scientific studies show that risk factors can interact to synergistically increase the likelihood of certain adverse and life-threatening outcomes. Yet, the extent to which individuals know that specific risk factor combinations present 'synergistic risks' is unclear and little is known about the determinants of such knowledge. This is largely because epistemological progress concerning this topic has been frustrated by a reliance on metrics that have latterly been judged to be of questionable validity. To address this issue, this paper presents two studies that assess an alternative approach (i.e. risk model judgements) which requires respondents to judge the risk for a factor combination relative to, rather than in isolation from, the risk attributable to each constituent factor. Results from both studies indicate that risk model judgements overcome the drawbacks of the traditional metrics. More importantly, the results provide epistemological insights into what can determine whether an individual understands that a factor combination presents a synergistic risk; these determinants include experiential and intuitive insights into the effects of combining specific risk factors, domain-specific judgemental experience and exposure to effective learning opportunities. These findings can be utilised in interventions aimed at helping individuals to make better decisions concerning multiple risk factors.

Keywords: Health; judgement; risk perception; synergistic risk

Running Head: Risk Model Judgements of Synergistic Risks

Adverse outcomes are rarely attributable to just one factor. For example, whilst some incidence of adverse cardiovascular events might be attributed to a single factor, such as obesity, it is more often the case that cardiovascular events result from a combination of two or more factors, such as smoking, poor diet, genetic predisposition, heavy alcohol consumption, sedentary lifestyle, and so on (Yusuf, Giles, Croft, Anda, & Casper, 1998). Similarly, whilst one could attribute climate change to the CO_2 emissions produced by human activity, the causal circumstances are far more complex. Specifically, factors, such as global population growth, affluence, technology, and consumption behaviours, all interact to determine the level of CO_2 emissions produced by humans (Huppert & Sparks, 2006; Satterthwaite, 2009; York, Rosa, & Dietz, 2003). Awareness of vulnerability to certain risk factors can play a major role in influencing the adoption of precautionary behaviours (Goodwin, Willson, & Stanley, 2005; Eiser & Arnold, 1999; Floyd, Prentice-Dunn, & Rogers, 2000) and it is, therefore, vital that individuals have the ability to understand the range of factors that might increase the risk of specific adverse events and, more importantly, how the interaction of these factors may increase the risk further.

The value of understanding the risk attributable to multiple factors is further accentuated by the accumulation of scientific evidence showing that certain factor combinations interact to present a risk that is greater than the sum of the risk attributable to each constituent factor. Such 'synergistic risks' have been reported across a range of psychological, sociological, ecological and epidemiological studies. For example, research has identified a synergistic risk of developing lung cancer for individuals who are regularly exposed to both radon (a naturally occurring radioactive gas) and tobacco smoke (Barros-Dios, Barreiro, Ruano-Ravina, & Figueiras, 2002; Darby et al., 2005; Pershagen et al., 1994). Furthermore, evidence indicates that the interaction of habitat destruction and climate change present synergistic risks of food shortages and the

extinction of various animal and plant species (Brook, Sodhi, & Bradshaw, 2008; Lin, Perfecto, & Vandermeer, 2008; Travis, 2003). The need to learn more about the extent to which individuals understand this important risk issue is highlighted by the sustained identification of factor combinations that present synergistic risks (e.g., Ben et al., 2011; Boltz, Hollenbeak, Ortenzi, & Dillon, 2012; Brook, Sodhi, & Bradshaw, 2008).

A clear concern is that if individuals underestimate the threat posed by synergistic risks they will not be motivated to adopt appropriate precautionary behaviours or to demand/support political, social, and economic action to address these issues (Dawson, Johnson, & Luke, 2012a, 2012c; French, Sutton, Kinmonth, & Marteau, 2006). Similarly, where policy-makers and regulators do not recognise that certain combined factors present synergistic risks, there is a concern that they may fail to implement policies and strategies that are proportionate to the risks (Berenbaum, 1989; Cogliano, 1997). The legitimacy of these concerns is confirmed by studies that have investigated individuals' risk estimates for combined factors that present synergistic health risks. Specifically, a majority of these studies found that individuals' estimates were consistent with either additive (i.e. equal to the sum of the constituent risks) or sub-additive (i.e. less than the sum of the constituent risks) models of risk (for a comprehensive taxonomy, see French et al., 2006). However, the authors of several studies exploring subjective judgements of synergistic risks have questioned the validity of the previous findings, arguing that the participants' risk magnitude judgements may have been artefacts of the linear rating scales employed (French, Gayton, Burton, Thorogood, & Marteau, 2002; French, Marteau, Sutton, & Kinmouth, 2004; French et al., 2006; Hampson, Andrews, Lee, Lichtenstein, & Barckley, 2000). For example, when participants have been required to use a Likert-type scale to provide separate risk magnitude estimates for constituent risk factors, it was observed that estimates were often at

the high end of the scale (e.g., a rating of '6' on a 7-point scale), which left no room on the scale to represent a synergistic effect for the factors when combined. Furthermore, French et al. (2006) found that risk estimates for the same combination of risk factors varied according to whether participants used a nine-point, 101-point, or unbounded response scale. Moreover, French et al.'s findings also indicated that the sensitivity of the three scales may not be sufficient to detect different risk models (i.e. sub-additive, additive, or synergistic) for different factor combinations (see French et al. [2004, 2006] for a thorough analysis and review of the validity of the metrics employed in previous studies). Consequently, few firm conclusions can be made from the studies that employed linear rating scales with regards to the extent that subjective judgements of synergistic risks are consistent with scientific risk assessments.

Using risk model judgements to assess judgements of synergistic risks

More recently, researchers have employed alternative methods, such as examining 'possibility judgements' and 'articulated reasoning' to assess individual's judgements and understanding of synergistic risks (see Dawson, Johnson, & Luke, 2012a, 2012c). One innovative approach, introduced by Condit and Shen (2011) in a study concerning public understanding of gene-environment interactions, required participants to provide risk *model* judgements rather than risk *magnitude* estimates. That is, instead of asking participants to provide single-point risk estimates on a linear continuum or scale, participants were asked to judge whether the risk attributable to the combination of (risky) health behaviours and a genetic predisposition to heart disease was either less than (i.e., sub-additive), equal to (i.e., additive) or more than (i.e. synergistic) the sum of the risk attributable to each factor alone. Condit and Shen reported that a majority of their participants judged that the combination presented an additive risk and that only one-third of the

participants judged, in line with scientific risk assessments, that the combination presented a synergistic risk. As discussed below, while these findings are mostly of specific value to those interested in lay understanding of harmful interactions between genes and lifestyle behaviours, the use of risk model judgements in research concerning individuals' understanding of synergistic risks has several potential advantages over the linear scales previously employed.

To improve the calibration of subjective judgements involving multiple probability estimates (i.e. risk judgements), an approach known as decomposition-recomposition can be employed (Goodwin & Wright, 2010, Hora, Dodd, & Hora, 1993). In this procedure, a judgement task involving multiple subjective probability estimates is broken-up (decomposed) into individual judgements and then recombined mechanistically (recomposed). The underlying assumption of this approach is that task complexity is reduced and, thus, the calibration of judges' estimates can be improved. However, the research evidence in this regard is mixed. Decomposition-recomposition has been shown to enhance the accuracy of probability judgements in some circumstances (Edwards, Phillips, Hays, & Goodman, 1968; Wright, Rowe, Bolger, & Gammack, 1994) but Wright, Saunders and Ayton (1988) found that this is not always the case. It has been proposed that the decomposition-recomposition approach may fail to improve probability judgements when the task is framed in a manner that is inappropriate to the judgement task or when it renders the task unfamiliar to the judge (Goodwin & Wright, 2010; Wright et al., 2009).

Previous studies investigating peoples' risk judgements for combined factors have obtained risk estimates from participants in a decomposed form. That is, participants have made separate (decomposed) risk estimates for each constituent factor and for the combined factors. The researchers then compared (mechanistically recomposed) the three risk estimates to

determine whether the individual's risk estimate for the combined factors is less than (sub-additive), equal to (additive) or more than (synergistic) the sum of the participant's risk estimates for each constituent factor. Whilst this approach may reduce task complexity, it may also fail to make the ultimate purpose of the judgement task explicitly clear to the participants, i.e., the need to make a risk judgement for the combined factors *relative* to the risk attributable to each constituent. Consequently, it is not explicitly clear to the respondent that he/she should articulate whether he/she believes that the combined factors presents a risk that is less than, equal to, or more than the sum of the risk presented by each constituent factor. Hence, even if a respondent is aware (consciously or intuitively) that a combination presents a synergistic risk, it cannot be assumed that he/she will appreciate that this should be reflected in the risk estimates that he/she provides. By contrast, the approach of asking participants to provide risk *model* judgements, as employed by Condit & Shen (2011), overcomes this issue because participants are explicitly instructed to consider whether the risk for the combined factors is less than, equal to or more than the sum of the risk for each constituent factor.

Research shows that when it is not made explicitly clear to participants exactly what beliefs/judgements they should represent in their responses to a risk judgement task, participants often respond in a way that is not intended by the researcher (for a review see Windschitl, 2002). For example, Fischhoff and Bruine de Bruin (1999) found, across multiple datasets, that there was a disproportionately high number of respondents who selected 50% on a probability scale, which Fischhoff and Bruine de Bruin identified as respondents using the 50% option to represent uncertainty (e.g. "I'm not sure; it's a 50-50 chance"). Also, Borland (1997) identified that respondents often used linear rating scales to represent subjective concern (i.e. the extent to which the individual is worried about a risk factor), rather than to provide probability estimates.

Consequently, the evidence indicates that when it is not made explicit to a participant what beliefs/judgements he/she should represent in his/her response, there is the possibility that they will provide data that reflects beliefs or judgements that differ from those under investigation.

A further reason why risk model (cf. risk magnitude) judgements may be advantageous in research concerning subjective judgements of synergistic risks is that individuals often have difficulty in accurately representing their perceptions of risk magnitudes in numerical/linear formats (Borland, 1997; Fischhoff & Bruine de Bruin, 1999; Windschitl, 2000). This may be because risk perceptions are not necessarily pre-formed as numerical probability estimates, but are often basic beliefs or fuzzy representations concerning the likelihood of potential threats (Brown & Morley, 2007; Eiser, 1994; Sjoberg, 2000). Also, research indicates that single-point risk estimates may not necessarily reflect an individual's more intuitive understanding of risk and uncertainty (Borland, 1997; Flugstad & Windschitl, 2003; Wallsten, Budescu, & Zwick, 1993; Windschitl, 2002; Windschitl & Wells, 1996). Hence, the task of representing risk perceptions as single-point risk estimates, both for single and combined risk factors, may either be unfamiliar or counterintuitive for many individuals; particularly, when it is unclear that the estimates should represent the effects of possible interactions. Thus, compared to the requirements of a linear rating scale, risk model judgements are advantageous because (a) less onus is placed on the respondent to be numerically-specific when providing a risk estimate and (b) the respondent can express a conceptual understanding of the directional effects of combining two specific risk factors in terms of a risk model,. Risk model judgements may, therefore, be more suited to enabling individuals to express an understanding of synergistic risk, particularly when that understanding is more tacitly, than numerically, encoded.

Examining the efficacy of risk model judgements in synergy studies

Whilst Condit and Shen's study has effectively illustrated how risk model judgements can be employed in research concerning subjective risk judgements for multiple factors, their application of this approach has a number of limitations. First, the approach was only used to assess judgements for one risk factor combination. Hence, although Condit and Shen reported that their results indicate the approach is "… *a psychometrically valid scale* …" (Condit & Shen, 2011, p. 115), this conclusion seems somewhat premature given that it is unclear whether the results were an artefact of the approach. That is, because the participants were not asked to provide risk model judgements for other combinations, it is not possible to determine whether the results are specific to the combination or specific to the response metric. Second, participants in Condit and Shen's study were asked to provide risk model judgements in response to a written scenario that used numerical data to describe the likelihood of heart disease for an individual who had both a genetic predisposition to the disease and unhealthy eating habits (e.g. the scenario stated that the gene increases the risk by 20%, and the eating habits increase the risk by 20%; participants were asked to judge whether the two factors operating together would result in a risk that was less than, equal to or more than 40%). As Condit and Shen identified (p. 122), the provision of numerical data may have framed the judgement task as a mathematical problem and, therefore, induced participants to employ numerically reasoning to 'solve the problem'. Previous research by Dawson, Johnson and Luke (2012a) shows that individuals who reason numerically when formulating a judgement about the risk attributable to combined factors often arrive at an additive risk model because they believe that the numerical risks 'add up' (i.e. 20% + 20% = 40%) consistent with the notion of 'adding' (i.e. combining) one risk factor to another. Hence, the use of a numerical frame by Condit and Shen could have influenced many participants to

reason mathematically and, consequently, to be biased towards adopting an additive risk model. Third, it is unclear from Condit and Shen's findings why different risk model judgements emerged within a relatively homogenous group of participants (i.e. approximately two-thirds of the sample indicated an additive risk model, and one-third indicated a synergistic risk model). That is, it is not known whether the different judgements arose due to different varieties of experience or knowledge within the group and, if so, what was the nature of this experience or knowledge. In other words, the study did not provide sufficient insight into the key question of why it is that some individuals do and others do not understand that certain factor combinations present synergistic risks. The two studies presented in this paper are aimed at addressing the concerns indicated above.

Several previous studies have assessed the validity of different psychometric scales employed to measure subjective risk judgements for combined hazards (e.g., French et al. 2002, 2004, 2006; Hampson, Andrews, Barckley, Lee & Lichtenstein, 2003). These assessments have typically been made via two approaches: First, the risk judgments obtained by the metric(s) are compared to anticipated results. For example, Hampson et al. (2003) predicted that the synergistic risk attributable to drinking-and-driving would be familiar to most people, and that a metric which obtained risk judgements reflecting knowledge of this synergy could be considered valid. Second, multiple metrics are employed, and the risk judgements obtained by each metric are compared; agreement between these judgements is interpreted as an indicator of validity. However, two distinct problems have emerged with this latter approach. First, there have been instances where two scales demonstrate agreement in one comparative study (indicating validity), but do not demonstrate agreement in another (see French et al., 2002, 2006). Second, where disagreement between different metrics is identified, there is no 'gold standard' measure

to which the results obtained from either metric can be compared for validation purposes. Consequently, in making our assessment of the validity of risk model judgements, we adopted the former approach and compared the results we obtained to predicted results. This approach can be described as an assessment of *criterion-related validity* (Litwin, 2003), and we adopted this method because it avoids the uncertainties, as described above, of assessing validity via comparisons between multiple metrics.

In Study 1, participants were asked to provide risk model judgements for one of three combinations that present synergistic health risks: alcohol-driving (= fatal collision), aspirin-clopidogrel (= internal bleeding), or radon-tobacco (= lung cancer). This design facilitated an assessment of whether risk model judgements would vary according to the factor combination under consideration and, therefore, whether the risk model approach is sufficiently sensitive to record such variations. In addition, in Study 1 the judgement task was presented to participants in either a numerical or verbal frame. This feature of the study design enabled an assessment of whether the framing of the risk model judgement task would influence participant's responses. Clearly, different responses between the numeric and verbal frames would suggest that the judgements were dependent on the description of the task at hand and, therefore, that the risk model metric may be low in face validity. However, it is important to note that our intention was not to address the broader question of whether numeric representations of risk influence or inhibit perceptions of synergistic risks in general. In Study 2, participants provided risk model judgements for the pharmaceutical combination of aspirin and clopidogrel. Participants consisted of one group of lay individuals and one group of domain experts (i.e. *Independent Prescribers*, such as doctors and pharmacists, who are qualified to prescribed drugs; referred to hereafter as 'IPs'). Importantly, the use of these two groups facilitated an assessment of whether risk model

judgements vary according to domain-specific knowledge and/or experience and, therefore, whether the risk model format was sufficiently sensitive to record this difference. Moreover, Studies 1 and 2 were designed to provide insights into the extent to which knowledge and/or experience of each risk factor combination might mediate whether an individual understands that the combination presents a synergistic risk.

STUDY 1

Method

Participants

One-hundred-and-six participants (49 men, 56 women, and 1 who did not indicate his/her gender) aged 18 to 75 (M = 35.79, SD = 17.69) were recruited online via five websites dedicated to social science research (e.g., www.onlinesocialpsychology.org, www.onlinepsychresearch.co.uk, etc.). Forty-three percent of participants were resident in the US, 40% in the UK and the remaining 17% in one of nine other Asian, Australasian, European or North American countries. Twenty-eight percent had completed a secondary/high school education, 27% had completed some postsecondary schooling and 44% had a university education. Participants were not rewarded for taking part in the study.

Design

The experiment followed a 3 (risk factor combination: alcohol and driving, radon and tobacco, aspirin and clopidogrel) x 2 (frame: numerical, verbal) design, with both factors as between-subjects. The dependent variable was the subjective risk model attributable to the factor combination (risk model: sub-additive, additive, or synergistic).

The three risk factor combinations presented in Study 1 have all been found, in practice, to present synergistic risks: alcohol-driving (Cherpitel, Tam, Midanik, Caetano, & Greenfield, 1995; Institute for Alcohol Studies, 2010; Office for National Statistics, 2009), aspirin-clopidogrel (Delaney, Opatrny, Brophy, & Suiss, 2007; Hallas et al., 2006) and radon-tobacco (Barros-Dios et al., 2002; Darby et al., 2005; Pershagen et al., 1994).[1] These combinations were selected to provide a set of synergistic risk combinations that were each likely to be of different degrees of familiarity to the participants and, therefore, could facilitate an assessment of the extent to which prior knowledge/experience of the combination would influence the participant's risk model judgements. Specifically, the synergistic risk of a fatal accident for the alcohol-driving combination would probably be familiar to most participants because of (a) widespread public awareness campaigns, (b) laws that prohibit drinking and driving (which implies that the factor combination is dangerous) and (c) the ease with which individuals can understand, whether as a result of direct or vicarious experience, that a serious accident is much more likely if the driver has a severely impaired ability to concentrate, judge distances, react quickly, etc. due to alcohol consumption. By contrast, the synergistic risk of gastro-intestinal bleeding presented by the aspirin-clopidogrel combination would probably not be familiar to most lay individuals and, therefore, it is unlikely that the study's participants would know much/anything about any potential risks and/or benefits of combining the two drugs. The extent to which the synergistic risk attributable to radon-tobacco would be familiar to the participants is unclear. More specifically, in contrast to the alcohol-driving combination, (a) the risk of combining radon and tobacco has not been highlighted by widespread public campaigns, (b) the combination is not prohibited in law and (c) the underlying mechanism for the radon-tobacco synergistic risk is unlikely to be understood by the participants, as the mechanistic process is complex and not yet

fully understood by scientists (Harley, Chittaporn, Heikkinen, Meyers, & Robbins, 2008). However, authors of both US- and UK-based studies, each employing different response metrics, have reported that participants judged the risk attributable to the radon-tobacco combination as synergistic (Eiser, Reicher, & Podpadec, 1995; Hampson et al., 2000).

Materials

The judgement task presented to each participant started with a short paragraph, comprised of two sentences, which participants read first (see Appendix). The first sentence stated that research evidence showed that the likelihood of an adverse outcome (e.g. gastro-intestinal bleeding) increased for an individual exposed to a specific single risk factor (e.g. taking aspirin). The second sentence stated that research evidence showed the likelihood of that same adverse outcome also increased for an individual exposed to a different single risk factor (e.g. taking the drug 'clopidogrel'). The wording employed in these two sentences was manipulated so that participants were presented with information that described the likelihood of the adverse outcome in either a 'verbal' or a 'numeric' frame. Specifically, participants in the verbal condition read sentences that described the likelihood as "*an increased chance*" for each factor. In contrast, participants in the numeric condition read sentences that described the likelihood for each factor in the form of a relative frequency, which had been derived from empirical research data (e.g. "*a 1 in 100,000 chance*"). Hence, participants in the verbal condition, unlike those in the numeric condition, were not provided with the 'objective' risk magnitude and, therefore, made their own subjective assessment of the risk attributable to each constituent factor. All relative frequencies (see Table 1) reported in the numeric condition featured the same

denominator to avoid the introduction of an unwanted framing effect (see Lipkus, 2007; Okan, Garcia-Retamero, Cokely, & Maldonado, 2011).

[Insert Table 1 about here]

After the first paragraph, participants were instructed to consider the likelihood of an individual experiencing the adverse outcome, as mentioned in the first paragraph (e.g. gastro-intestinal bleeding), when exposed to both factors (e.g., aspirin *and* clopidogrel). Participants were then asked to state, using a multiple-choice response format, whether they judged this likelihood to be "*less than*", "*equal to*" or "*more than*" the likelihood for an individual exposed to only the first factor (e.g. aspirin) "*added to*" the likelihood for an individual exposed to only the second factor (e.g. clopidogrel). The layout, style and content of the judgment task were developed in accordance with guidelines outlined by Osterlind (1998) for the construction of multiple-choice response items.

Having completed the judgement task, participants then read the following instruction "*Please indicate how confident you are that your judgment (i.e., less than, equal to, or more than) in the previous task represents what has been found in scientific research, where: 0% = "I have no idea whether my judgment represents what has been found in scientific research" and 100% = "I am certain my judgment represents what has been found in scientific research"*." To respond to this question, participants could select one of eleven categorical options distributed in ten percent intervals (i.e. 0%, 10%, 20% … 100%). The purpose of this question was to measure the extent to which participant's believed that their risk model judgements were veridical (i.e. consistent with whether the 'objective' evidence indicates the risk model is sub-additive, additive or synergistic) and, therefore, whether participants believed they were drawing on knowledge of objective data to formulate a veridical subjective judgement. In addition,

participants were asked to state whether, prior to participating in the study, they were aware that (a) aspirin increases the risk of gastrointestinal bleeding, (b) clopidogrel increases the risk of gastrointestinal bleeding and (c) radon exposure increases the risk of lung cancer (it was assumed that participants would be aware that either alcohol use or vehicle driving increases the risk of an accident and that tobacco smoking increases the risk of lung cancer).

Procedure

Each participant was presented with the experimental materials via an online survey system. The system randomly allocated participants to receive one judgement task concerning either the alcohol-driving, radon-tobacco or aspirin-clopidogrel combinations. The system also randomly allocated participants to receive the judgement task in either the numerical or verbal frame. Each 'factor combination group' consisted of 17 or 18 participants who received the task in a verbal frame and 17 or 18 who received the same task in the numeric frame. Participants were advised that they had an unlimited time to participate in the study (M = 10 minutes 28 seconds), to complete the questions in a place where they would not be distracted and not to consult any materials or persons whilst participating.[2]

Statistical analysis

The dependent and independent variables in Study 1 were, primarily, categorical variables. In order to test for main effects and interactions between these variables, we analysed the data using hierarchical loglinear analysis. Significant interactions were further analysed using chi-square (χ^2) tests. The data concerning the extent to which each participant believed his/hers risk

judgements to be veridical was treated as a continuous variable, and was analysed using ANOVAs.

Results

Risk judgements

Categorical risk judgements for each of the three combinations are displayed in Figure 1. An initial visual inspection of this descriptive data indicated that a majority of participants judged that the alcohol-driving and tobacco-radon combinations would present a synergistic risk (77% and 74%, respectively). However, the proportion of participants who judged that the aspirin-clopidogrel combination would present a synergistic risk (42%) was approximately equal to the proportion who judged that the combination would present an additive risk. These observations presented the first indication that the metric employed was sufficiently sensitive to record variations in risk model judgements for different risk factor combinations.

[Insert Figure 1 about here]

Before the hierarchical loglinear analysis was performed, an assessment was made to determine whether the data met the test assumption that no less than 20% of the data cells should have expected frequencies less than 5 (Field, 2009). However, because few participants judged the risk attributable to the combined factors as 'sub-additive', the data did not meet this assumption for three-way analyses (77% of cells featured expected frequencies less that 5). Consequently, the 'additive' and 'sub-additive' categories were collapsed into one category, which resulted in two categories for the risk model variable (i.e. 'synergistic' and 'non-synergistic'). Collapsing categories in this way is common in research that employs hierarchical loglinear analysis (e.g. Chung, 1996; Fairclough, Boddy, Hackett, & Stratton, 2009), and is

appropriate in circumstances, such as those here, where the research objectives (i.e. investigating whether or not participants judged the risk attributable to the factor combinations as a synergistic risk) are not impeded (Field, 2009). Collapsing the categories resulted in data that met the test assumptions and made the three-way analysis viable.

Risk judgements, risk factor combination, and framing

To assess whether participants' judgements varied according to factor combination and/or task frame, a three-way hierarchical loglinear analysis was performed: risk model x factor combination x frame. This identified a significant main effect for risk model, χ^2 (1) = 8.61, p < 0.01, that was qualified by a significant two-way interaction between the factor combination and risk model, χ^2 (2) = 11.91, $p < 0.01$. The main effect for risk model was attributable to the greater proportion of factor combinations overall that were judged to present a synergistic risk (64%) rather than a non-synergistic risk (36%), χ^2 (1) = 8.49, $p < 0.01$. The two-way interaction between factor combination and risk model was investigated via separate analysis of participants' risk judgements for each factor combination. This revealed that a significantly greater proportion of participants judged that the alcohol-driving combination would present a synergistic (cf. non-synergistic) risk, χ^2 (1) = 10.31, $p < 0.001$. Similarly, a significant majority of participants in the radon-tobacco group judged that the combination would present a synergistic (cf. non-synergistic) risk, χ^2 (1) = 8.26, $p < 0.01$. However, there was no significant difference between the proportion of participants who judged that the aspirin-clopidogrel combination would present a synergistic (42%) and a non-synergistic risk (58%), χ^2 (1) = 1.00, p = 0.32. No other significant interactions or main effects were found, χ^2s (1) <= 0.73, ps > 0.39. Hence, the results showed that the participants' risk judgements did not differ significantly

whether they were presented with the judgement task in a numeric or verbal frame. This was the case irrespective of the factor combination under consideration.[3]

Confidence in veridicality of risk judgements

A 3 (factor combination) x 2 (frame) independent measures ANOVA was performed on the veridicality judgement data. This identified that the differences in veridicality judgements between the aspirin-clopidogrel (M = 53.61, SD = 29.19), radon-tobacco (M = 61.14, SD = 32.88) and alcohol-driving (M = 67.35, SD = 29.57) groups were not significant, $F(2, 99) = 1.72$, $p = 0.19$. Furthermore, the difference in veridicality judgements between the numeric (M = 61.92, SD = 31.00) and verbal (M = 59.25, SD = 30.88) frame conditions were also non-significant, $F(1, 99) = 0.19$, $p = 0.66$. Although a significant interaction was identified between factor combination and frame, $F(2, 99) = 4.97$, $p < 0.01$, simple effects tests (employing a Bonferroni correction for multiple comparisons) for each factor combination indicated that the differences between judgements in each framing condition were not significant, $ts(18) <= 2.36$; $ps > 0.028$. Specifically, for the aspirin-clopidogrel combination, judgements in the numeric frame condition (M = 62.22, SD = 24.15) were higher than in the verbal frame condition (M = 45.00, SD = 31.86); similarly, for the radon-tobacco combination, judgements were higher in the numeric condition (M = 67.78, SD = 31.54) than the verbal condition (M = 54.12, SD = 33.74); however, for the alcohol-driving combination judgements were lower in the numeric condition (M = 55.00, SD = 37.24) than the verbal condition (M = 78.33, SD = 14.25).

Discussion

The results of Study 1 indicate that the risk model judgement format is sufficiently sensitive to record different risk judgements for different combinations. In other words, the judgements do not appear to be purely an artefact of the 'risk model judgement' metric. Furthermore, participants' risk model judgements did not differ as a result of manipulations in the framing of the judgement task (i.e. numerical vs. verbal). That is, the provision of either numerical or verbal risk magnitude information did not lead to a disproportionate increase in one specific risk model judgements. Hence, the results offer some evidence towards the format's face validity, as the translation of the risk model construct did not vary according to the task frame.

In assessing the results of Study 1 from an epistemological perspective, it is pertinent to consider the mediating effect of the familiarity of the risk factor combination. Certainly, the familiarity of the potentially adverse effects of combining alcohol and driving provides a sound explanation for the finding that a large majority of participants understood that the combination presents a synergistic risk. Similarly, the unfamiliarity of the potential adverse effects of combining aspirin and clopidogrel (only 14% of participants reported that prior to the study they were aware that separate use of clopidogrel can lead to gastrointestinal bleeding) provides a plausible explanation for the finding that less half the participants understood that the combination presents a synergistic risk. From this perspective, the finding that a large majority of the participants judged that the radon-tobacco combination presents a synergistic risk would suggest that the harmful effects of this combination were familiar to many of the participants. However, further analysis did not support this interpretation. That is, within the radon-tobacco condition, the risk model judgements of participants who reported that they were/were not previously aware that exposure to radon increases the risk of lung cancer (n = 20/15) did not

differ significantly, χ^2 (2) = 0.45, p = 0.50. This then raises the question of why a majority of the participants in Study 1 judged that the radon-tobacco combination presents a synergistic risk.

Whilst it is highly unlikely that the participants would have formulated a detailed mental model of the complex bio-chemical interaction between radon and tobacco, the participants may have constructed a more rudimentary mental model that encapsulated the vulnerability of the human body to the potent effects of simultaneous expose to tobacco toxins and radiation. Such an intuitive understanding of synergistic risk has been identified in previous research (see Dawson, Johnson & Luke, 2012a) and, therefore, might also offer a partial explanation for other evidence indicating that individuals may hold synergistic models of risk for the radon-tobacco combination (Eiser et al., 1995; Hampson et al., 2000). However, this interpretation raises the additional question of why a majority of participants did not employ this intuitive synergistic risk model when making a judgement for the aspirin-clopidogrel combination? A possible explanation is that, as indicated earlier, most individuals probably know very little about the effects of combining aspirin and clopidogrel and, more specifically, whether these two drugs might also result in health benefits as well as risks. Notably, research shows that lay individuals tend to perceive most medicinal drugs as beneficial (Kraus, Malmfors, & Slovic, 1992), and that subjective judgements' of risks and benefits are often confounded in individuals' minds (Alhakami & Slovic, 1994; Finucane, Alhakami, Slovic, & Johnson, 2000). Consequently, it is possible that many of the participants in Study 1 believed that the overall risk for the aspirin-clopidogrel combination may be attenuated by the potential health benefits of the combination. This may have led them to perceive the two drugs as a safer combination, rather than a combination that would present a synergistic risk. The finding that the highest number of non-

synergistic risk judgements was observed in the aspirin-clopidogrel condition provides evidence in support of this interpretation.

Taken together, the results of Study 1 indicate that (a) prior knowledge of (i.e. familiarity with) the harmful effects of combining two risk factors can aid judgements that the combination presents a synergistic risk, yet (b) prior knowledge of the harmful effects of combining certain factors is not necessarily a prerequisite to arrive at a judgement that a risk factor combination (e.g. radon and smoking) presents a synergistic risk. The results also show that participants' confidence about the calibration of their subjective judgements with 'objective' scientific risk assessments did not vary significantly between the three 'factor combination' conditions. This suggests that awareness of 'objective' data may not be a prerequisite for understanding that specific factor combinations present synergistic risks and that many individuals may base their risk model judgements on experiential knowledge and/or intuitive cognitions. Again, this could explain why many participants attributed a synergistic risk model to the radon-tobacco combination irrespective of whether they had some prior knowledge of the scientific literature concerning the risk attributable to the factors.

STUDY 2

If the absence of pharmacological/domain-specific knowledge concerning the risk/benefits of the aspirin-clopidogrel combination was a factor that influenced judgements in Study 1, then one might expect that individuals with such knowledge (i.e. IPs) would judge that a synergistic risk is attributable to the combination. Similarly, one might also expect IPs to be more confident regarding the extent to which their risk judgements for the combination are veridical. To investigate these predictions, a second study was conducted in which a group of IPs and a group

of lay individuals were asked to judge whether the risk attributable to the combination of aspirin-clopidogrel was less than, equal to or more than the sum of the risk for each of the constituent drugs. Importantly, the data obtained in this study facilitated an assessment of the discriminant validity of the risk model task format. That is, a significant difference between the responses of the two groups would provide evidence of the formats capacity to distinguish between different risk model judgements for different populations (Laver-Fawcett, 2007).

Method

Participants

To recruit a sample of IPs and a sample of non-experts matched to the IPs on key socio-demographics (i.e., age, gender, education, and national residence), chain-referral sampling was employed (Penrod, Preston, Cain, & Starks, 2003). We initially made contact with individuals (known as 'gatekeepers') who, by virtue of their employment, were in a position to recruit other individuals that met the socio-demographic profile suitable for participation in the study (Penrod et al., 2003). Several gatekeepers were employed in the recruitment of both IPs and non-experts to minimise selection bias (Atkinson & Flint, 2003), and the gatekeepers did not participate in the study. Participants and gatekeepers were not rewarded for taking part in the study.

The sample of IPs (n = 31) were recruited by five gatekeepers, who were either participants in a university alumni network, mediated an online forum for pharmacists or chaired a professional association for UK pharmacists. The recruited participants consisted of 17 men and 13 women (one participant's gender not stated), and their ages ranged from 28 to 59 years (M = 42.40, SD = 9.97). We established, via a questionnaire, that the 31 participants met all of the following criteria: resided in the UK, had qualified to independently prescribe prescription-

only drugs, were employed in a role where they independently prescribed drugs and considered the potential for adverse drug-drug interactions at least once per week in their professional role.

Our sample of non-experts (n = 30) were recruited by six gatekeepers from a range of UK-based public and private sector organisations. The sample of participants held a variety of professional positions that included Head Teacher, Project Manager, Financial Analyst, Administrator and Chartered Accountant. The non-expert participants were matched to the IPs in terms of being UK residents, aged between 26 and 64 (M = 39.73, SD = 10.14) and educated to Bachelor's Degree level or higher. Consistent with the IP group, the non-expert group consisted of 14 men and 16 women. There was no significant difference between the mean age of the IP and non-expert groups, $t(58) = 1.03$, $p = 0.31$, and there was no significant difference between the proportion of men/women in the two groups, χ^2 (1) = 0.60, $p = 0.27$. Questionnaire-elicited data was obtained from the non-expert participants to ensure none of them met the criteria for inclusion in the IP group (see above).

Design

The experiment followed a 2 (domain expertise: IP, non-expert) x 2 (frame: numerical, verbal) design with both factors as between-subjects. The dependent variable was the subjective risk model attributable to the factor combination (risk model: sub-additive, additive, or synergistic).

Materials

The design of the judgement task materials presented to each participant replicated that employed in Study 1, but in Study 2, only the aspirin-clopidogrel judgement task was presented. After completing the task, participants provided veridicality judgements (as per Study 1) and

were asked to state whether, prior to participation in the study, they were aware that (a) use of aspirin increases the risk of gastro-intestinal bleeding and (b) use of clopidogrel increases the risk of gastro-intestinal bleeding.

Procedure

The experimental materials were presented to all participants via an online survey system. The system randomly allocated the judgement task in either the numerical or verbal frame within both the non-expert and IP groups. Participants received the same instructions provided in Study 1. The mean participation time was 10 minutes 47 seconds.

Statistical analysis

Data analysis was performed as per Study 1. As observed in Study 1, the proportion of cells with an expected count of 5 or more was not sufficient to meet the assumption for a three-way hierarchical loglinear analysis (58% of cells featuring expected frequencies less that 5). To overcome this issue, the 'additive' and 'sub-additive' categories were collapsed into one category (i.e., 'non-synergistic'). This resulted in data that met the test assumptions, making the three-way analysis viable.

Results

Risk judgements

Categorical risk judgements for both the non-expert and IP groups are displayed in Figure 2. An initial visual inspection of this descriptive data indicated that a majority (68%) of the IPs judged that the combination would present a synergistic risk. In contrast, the proportion of non-experts

who judged that the combination would present a synergistic risk (37%) was equal to the proportion who judged that the combination would present an additive risk. These observations presented the first indication that the risk model format possessed discriminant validity.

In the following inferential analysis, the 'additive' and 'sub-additive' categories were collapsed into one category.

[Insert Figure 2 about here]

Risk judgements, expertise, and framing

To assess whether participants' risk judgements varied according to expertise and/or task frame, a three-way hierarchical loglinear analysis was performed: risk model x domain expertise x frame. This identified a significant two-way interaction between expertise and risk model, χ^2 (1) = 5.99, $p < 0.05$. No other significant interactions or main effects were found, χ^2s (1) <= 0.41, *p*s > 0.52. Hence, participants' risk judgements did not differ significantly between the numeric or verbal frame conditions and this was the case across both the IP and non-expert groups.

The two-way interaction between expertise and risk model was investigated via separate analyses of risk judgements for each expertise group. This revealed that a significantly greater proportion of IPs judged that the aspirin-clopidogrel combination would present a synergistic, rather than non-synergistic, risk, χ^2 (1) = 3.90, $p < 0.05$. However, there was no significant difference between the proportion of non-experts who judged that this combination would present a synergistic risk and those who judged that the combination would not, χ^2 (1) = 2.13, $p = 0.14$.

Confidence in veridicality of risk judgements

A 2 (domain expertise) x 2 (frame) independent measures ANOVA was performed on the veridicality judgement data. This identified a significant main effect for expertise, $F(2, 57) = 3.14$, $p = 0.04$ (employing a lopsided test of significance; Abelson, 1995; Levine & Banas, 2002), with IPs being significantly more confident ($M = 57.10$, $SD = 32.48$) than non-experts ($M = 42.67$, $SD = 30.00$) regarding the extent to which their risk model judgements were veridical. However, no other significant main effects or interactions were identified, Fs$(1, 57) > 0.47$, ps < 0.40.

Prior knowledge of constituent risk factors

Ninety-four-percent of IPs reported that prior to participation in the study he/she was aware that separate use of either aspirin *or* clopidogrel increases the risk of gastro-intestinal bleeding. By contrast, only ten percent of non-experts reported being aware of this side-effect for separate use of both drugs.

Discussion

Importantly, the results of Study 2 demonstrate that the format can help to distinguish between different response patterns for different groups; thus, providing evidence of the format's discriminant validity. Also, the similarity between the judgement pattern observed for the aspirin-clopidogrel combination in Study 1 and for the non-experts in Study 2 offers some evidence of the format's external validity. Moreover, Study 2 revealed no effect for the verbal/numeric framing manipulations. This provides further evidence that the risk model judgement task provides a valid means of assessing risk judgements for combined factors.

Previous studies examining the risk judgements of experts indicate that experts are more likely to show "good" judgemental performance when the *ecological validity* (i.e. the degree to which the expert makes a judgement within their professional domain) and the *learnability* (i.e. the degree to which the experts' judgemental veridicality has been improved by the availability of objective data or usable feedback) of the judgement task are high (Bolger & Wright, 1994; Rowe & Wright, 2001). The evidence from Study 2 suggests that the judgement task was more ecologically valid for the IPs because this group of participants (a) regularly considered the effects of drug-drug interactions in their work role and (b) were much more aware of the side-effects attributable to both aspirin and clopidogrel when taken independently. The results also suggest that learnability probably played a role in the IPs' risk model judgements, because this group of participants was more confident that their judgements were consistent with objective risk assessments for the aspirin-clopidogrel combination. Hence, the judgement task presented to participants in Study 2 appears to have been one that was higher in ecological validity and learnability for participants in the IP group. It, therefore, seems reasonable to assert that these two factors played a role in the judgements made by the IPs, who demonstrated greater veridicality (cf. non-experts) in their risk judgements for the factor combination of aspirin-clopidogrel.

General Discussion

The two studies presented here provide evidence supporting the validity of using risk model judgements to assess whether individuals understand that a particular combination of factors presents a synergistic risk. Furthermore, the results of the two studies also provide important insights into some of the factors that may mediate such an understanding. Importantly, the results

of Study 1 showed that many lay individuals can/do make veridical risk judgements for factor combinations that have been found, in practice, to present synergistic risks. However, it is of concern that both Studies 1 and 2 also found that, for lay individuals, this veridicality does not necessarily extend to all factor combinations. Reassuringly, the greater proportion of veridical risk model judgements made by the domain-experts in Study 2 indicates that knowledge of the synergistic risk attributable to certain combinations can be learned. This suggests that laypersons should also be able to acquire a more veridical understanding of the synergistic risk attributable to certain factor combinations, provided they are exposed to effective learning opportunities (for supporting evidence see Dawson, Johnson & Luke, 2012c).

One possible interpretation of the results of Study 1 could be that perceptions of synergistic risks are inversely proportionate to the ratio of the difference of risk magnitude between each constituent risk factor (the drug interaction ratio is 10:1, radon tobacco 17:4 and alcohol-driving 3:1). That is, individuals may be less likely to arrive at a synergistic model of risk when the constituent risk factors differ greatly in risk magnitude. However, the results of both our studies highlight two reasons to question this possible interpretation. First, in both studies the exact details of the risk magnitudes were provided to the participants in the 'numeric frame' conditions (hence, these participants could have used this information to identify the ratios). Yet, the results showed that the judgements made by these "informed" participants did not differ significantly from those of the "uninformed" participants in the 'verbal' frame condition (who were only advised that each risk factor presented "*an increased chance*" of the relevant adverse outcome). Second, it seems reasonable to infer that, due to their education and experience, the group of 'domain experts' in Study 2 would be more likely than the group of 'non-experts' to possess veridical knowledge of the separate risk magnitudes attributable to each

of the two drugs (aspirin and clopidogrel). Therefore, these experts would be more likely to possess a veridical understanding of the ratio of difference in the risk magnitudes of each drug (i.e., 10:1). However, the results of Study 2 show that it was the group of non-experts (cf. 'experts') who were less likely to arrive at a synergistic risk model for the aspirin-clopidogrel combination. Hence, the evidence from both studies indicates that knowledge of the specific risk magnitudes attributable to each constituent factor (and, therefore, knowledge of the difference of the ratio between the constituent risks) does not influence the risk judgments for the combination of factors.

The flip-side of the findings from Study 2 is that nearly a third of the IPs did not demonstrate an awareness of the synergistic risk attributable to the aspirin-clopidogrel combination. Hence, consideration should also be given to how this lack of awareness might lead IPs to prescribe hazardous drug combinations and, therefore, how improving the accuracy of such risk judgements amongst IPs could lead to improvements in poly-pharmaceutical patient care. Extrapolating further, the non-veridical risk model judgements by some IPs suggests that it would be wrong to assume that expertise, in any domain, invariably leads to accurate risk model judgements for combined factors.

The data obtained in both studies reflects participants' risk *model* judgements, and not specific risk *magnitude* estimates. Hence, one could argue that, even where a person judges that a combination presents a synergistic risk, the individual may still not have an accurate understanding of the combination's specific risk magnitude and that this might impede sound risk-related decision making. However, it may be too much to expect most non-experts to make precise estimates of synergistic risk magnitudes, because this is a complex task requiring an understanding of the way such magnitudes vary in relation to the intensity, duration, proximity

and frequency of exposure to each constituent factor (Berenbaum, 1989; Cogliano, 1997). It, therefore, seems more pragmatic for researchers to focus on assessing whether individuals understand that a specific factor combinations present a synergistic model of risk, rather than to attempt to assess whether they are aware of the exact extent to which the combined risk exceeds the sum of the constituent risks (for a similar assertion see Hampson et al., 2003, p.1029). Hence, there is clearly a case for assessing risk model (cf. magnitude) judgements and, therefore, for employing the metric successfully utilised in the studies presented here.

Limitations and future directions

There are some limitations to the present research that could be addressed by future studies. First, only risk judgements for synergistic risk combinations were assessed and, therefore, the extent to which individuals' judgements for sub-additive or additive risk combinations are veridical was not assessed. Similarly, the extent of the validity and reliability of the risk model metric may only be firmly established via a greater number and variety of applications. Hence, future studies should use the risk model metric under controlled conditions to assess judgements for a wide range of sub-additive, additive and synergistic risk combinations, in relation to a variety of domains (e.g. health, ecological, social, etc.), and with respondents of varying degrees of domain expertise. Second, the comparison of the judgements of non-experts and domain-experts focused on only one drug-drug combination and, consequently, the differences identified in the study may or may not be evident for judgements of other drug combinations. Third, the use of internet-based samples provided many advantages (e.g. access to hard-to-reach populations, a diverse sample, etc.; see Reips, 2000) but limited the ability to control the conditions under which each participant operated; hence, some unknown biases may exist in the

data (e.g. multiple participation by one individual). Fourth, whilst our results indicate that risk model judgements do not vary according to the task frame (i.e., numeric vs. non-numeric), we suggest that caution is exercised in assuming that (a) variations in other aspects of the task frame will not influence judgments or (b) task frames will not influence judgments when other response metrics are employed. Fifth, we assessed the validity of the risk model metric by comparing the participants' judgements to anticipated results. We recognise that the validity of our findings largely rests on the extent to which the original rationale(s) for the anticipated results are deemed to be reliable and accurate. Future research in this field could aim to increase the robustness of the rationale(s) behind such anticipated results. For example, literature-informed rationales such as those utilised in our studies could be bolstered, for instance, by conducting pre-test interviews with individuals from the target population to explore the extent to which they demonstrate the anticipated awareness of the synergistic risks under investigation. Finally, our studies employed three risk factor combinations that varied in terms of (a) the magnitude of risk attributable to each of the constituent factors and (b) the ratio of difference in the risk magnitude attributable to each constituent factor within each combination. Future research could specifically examine whether either of these two variables influence an individual's risk model judgements for combined risk factors. For example, research could explore the possibility that individuals may be more inclined to arrive at a synergistic risk model judgement when both constituent factors present a relatively 'high' perceived risk, and whether individuals are less inclined to arrive at a synergistic model of risk when either one or both constituent factors is perceived to present a relatively 'low' or 'negligible' risk. To our knowledge, there is an absence of empirical studies that specifically assess the potential influence of these two variables.

Conclusion

The accumulation of evidence showing that various factor combinations present synergistic risks highlights a pressing need to learn more about individuals' understanding of this issue. Based on the results of the two studies reported here, the risk model judgement metric has demonstrated that it can serve as a useful tool for addressing this need. Furthermore, Study 2 is, to the best of our knowledge, the first to assess domain-experts' risk judgements for a factor combination that presents a synergistic risk. These studies have helped to identify a number of possible explanations as to why some individuals make veridical judgements of synergistic risks and others do not. These include familiarity with the harmful effects of combining the factors, and the extent to which the judge possesses both domain-specific knowledge and judgemental experience concerning the target combination. Consequently, we recommend that future research explores the efficacy of interventions/communications that aim to develop these characteristics in individuals. However, a vital tool for assessing the relative effectiveness of such interventions/communications is a valid method of measuring the recipients' risk judgements for combined factors. The results presented here demonstrate that the risk model judgement metric has the potential to fulfil this important task.

References

Abelson, R. P. (1995). *Statistics as Principled Argument*. Hillsdale, NJ: Erlbaum.

Alhakami, A. S. S., P. (1994). A Psychological Study of the Inverse Relationship between Perceived Risk and Perceived Benefit. *Risk Analysis, 14*(6), 1085-1096.

Atkinson, R., & Flint, J. (2003). Sampling, snowball: Accessing hidden and hard-to-reach populations. In R. L. Miller, & Brewer, J.D. (Ed.), *The A-Z of Social Research: A Dictionary of Key Social Science Research Concepts* (pp. 274-280). London: Sage.

Barros-Dios, J. M., Barreiro, M.A., Ruano-Ravina, A., Figueiras, A. (2002). Exposure to residential radon and lung cancer in Spain: A population-based case-control study. *American Journal of Epidemiology, 156*(6), 548 - 555.

Ben, Q., Cai, Q., Li, Z., Yuan, Y., Ning, X., Deng, S., & Wang, K. (2011). The relationship between new-onset diabetes mellitus and pancreatic cancer risk: A case-control study. *European Journal of Cancer, 47*(2), 248-254.

Berenbaum, M. C. (1989). What is Synergy? *Pharmacological Reviews, 41*(2), 93-141.

Bolger, F. W., G. (1994). Assessing the quality of expert judgment: Issues and analysis. *Decision Support Systems, 11*, 1-24.

Boltz, M. M., Hollenbeak, C. S., Ortenzi, G., & Dillon, P. W. (2012). Synergistic Implications of Multiple Postoperative Outcomes. *American Journal of Medical Quality*. doi: 10.1177/1062860611429612

Borland, R. (1997). What do people's estimates of smoking related risk mean? *Psychology and Health, 12*, 513-521.

Brook, B. W., Sodhi, N. S., & Bradshaw, C. J. A. (2008). Synergies among extinction drivers under global change. *Trends in Ecology & Evolution, 23*(8), 453-460. doi:

10.1016/j.tree.2008.03.011

Brown, S. L., & Morley, A.M. (2007). Risk perception, fuzzy representations and comparative optimism. *British Journal of Psychology, 98*(4), 575-587.

Cherpitel, C. J., Tam, T., Midanik, L., Caetano, R., & Greenfield, T. (1995). Alcohol and non-fatal injury in the U.S. general population: A risk function analysis. *Accident Analysis and Prevention, 27*(5), 651-661.

Chung, S. (1996). Performance effects of cooperative strategies among investment banking firms: A loglinear analysis of organizational exchange networks. *Social Networks, 18*(2), 121-148.

Cogliano, V. J. (1997). Plausible upper bounds: Are their sums plausible? *Risk Analysis, 17*(1), 77-84.

Condit, C. M., & Shen, L. (2011). Public Understanding of Risks from Gene-Environment Interaction in Common Diseases: Implications for Public Communications. *Public Health Genomics, 14*(2), 115-124.

Darby, S., Hill, D., Auvinen, A., Barros-Dios, J.M., Baysson, H., Bochicchio, F., Deo, H., Falk, R. Forastiere, F., Hakama, M., Heid, I., Kreienbrock, L., Kreuzer, M., Lagarde, F., Mäkeläinen, I., Muirhead, C., Oberaigner, W., Pershagen, G., Ruano-Ravina, A., Ruosteenoja, E., Schaffrath Rosario, A., Tirmarche, M., Tomáscaronek, L., Whitley, E., Wichmann, H-E., & Doll, R. (2005). Radon in homes and risk of lung cancer: Collaborative analysis of individual data from 13 European case-control studies. *British Medical Journal, 330,* 223-228.

Dawson, I. G. J., Johnson, J. E. V., & Luke, M. A. (2012a). Do People Believe Combined Hazards Can Present Synergistic Risks? *Risk Analysis*, *32*(5), 801-815. doi:

10.1111/j.1539-6924.2011.01693.x

Dawson, I. G. J., Johnson, J. E. V., & Luke, M. A. (2012b). Subjective judgements of synergistic risks: A cognitive reasoning perspective. *British Journal of Psychology, 103*(2), 203-223. doi: 10.1111/j.2044-8295.2011.02065.x

Dawson, I. G. J., Johnson, J. E. V., & Luke, M. A. (2012c). Helping Individuals to Understand Synergistic Risks: An Assessment of Message Contents Depicting Mechanistic and Probabilistic Concepts. *Risk Analysis*, no-no. doi: 10.1111/j.1539-6924.2012.01878.x

Delaney, J. A., Opatrny, L., Brophy, J.M., & Suissa, S. (2007). Drug–drug interactions between antithrombotic medications and the risk of gastrointestinal bleeding. *Canadian Medical Association Journal, 117*(4), 347-351.

Edwards, W., Phillips, L.D., Hays, W.L., & Goodman, B.C. (1968). Probabilistic information processing systems. *IEEE Transactions on System Science and Cybernetics, 4*, 248-265.

Eiser, J. R. (1994). Risk judgments reflect belief strength, not bias. *Psychology and Health, 9*(3), 197-199.

Eiser, J. R., Reicher, S.D. & Podpadec, T.J. (1995). Smokers' and non-smokers' estimates of their personal risk of cancer and of the incremental risk attributable to cigarette smoking. *Addiction Research, 3*, 221-229.

Eiser, J. R. A., B.W.A. (1999). Out in the midday sun: Risk behaviour and optimistic beliefs among residents and visitors on Tenerife. *Psychology and Health, 14*(3), 529-544.

Fairclough, S. J., Boddy, L.M., Hackett, A.F., Stratton, G. (2009). Associations between children's socioeconomic status, weight status, and sex, with screen-based sedentary behaviours and sport participation. *International Journal of Pediatric Obesity, 4*(4), 299-305.

Field, A. (2009). *Discovering Statistics Using SPSS* (3rd ed.). London: SAGE Publications.

Finucane, M. L., Alhakami, A., Slovic, P. & Johnson, S.M. (2000). The Affect Heuristic in Judgment of Risk and Benefits. *Journal of Behavioral Decision Making, 12*, 1-17.

Fischhoff, B., & Bruine de Bruin, W. (1999). Fifty-fifty = 50%? *Journal of Behavioral Decision Making, 12*(2), 149-163.

Floyd, D. L., Prentice-Dunn, S. & Rogers, R.W. (2000). A meta-analysis of research on protection motivation theory. *Journal of Applied Social Psychology, 30*, 407-429.

Flugstad, A. R., & Windschitl, P. D. (2003). The influence of reasons on interpretation of probability forecasts. *Journal of Behavioral Decision Making, 16*(2), 107-126.

French, D. P., Gayton, E.L., Burton, J., Thorogood, M. & Marteau, T.M. (2002). Measuring perceptions of synergistic circulatory disease risk due to smoking and the oral contraceptive pill. *Risk Analysis, 22*, 1139-1151.

French, D. P., Marteau, T.M., Sutton, S. & Kinmouth, L. (2004). Different measures of risk perception yield different patterns of interaction for combinations of hazards: smoking, family history and cardiac events. *Journal of Behavioral Decision Making, 17*(5), 381.

French, D. P., Sutton, S., Kinmonth, A.L., & Marteau, T.M. (2006). Assessing perceptions of risks due to multiple hazards. *Journal of Risk Research, 9*(6), 657-682.

Goodwin, R., Willson, M., & Stanley, G. (2005). Terror threat perception and its consequences in contemporary Britain. *British Journal of Psychology, 96*(4), 389-406. doi: 10.1348/000712605x62786

Goodwin, P., & Wright, G. (2010). The limits of forecasting methods in anticipating rare events. *Technological Forecasting and Social Change, 77*(3), 355-368.

Hallas, J., Dall, M., Andries, A., Sogaard-Andersen, B., Aalykke, C., Moller-Hansen, J.,

Andersen, M., & Touborg-Lassen, A. (2006). Use of single and combined antithrombotic therapy and risk of serious upper gastrointestinal bleeding: population based case-control study. *British Medical Journal, 333*, 726-730.

Hampson, S. E., Andrews, J.A., Lee, M.E., Lichtenstein, E. & Barckley, M. (2000). Radon and cigarette smoking: Perceptions of this synergistic risk. *Health Psychology, 19*, 247-252.

Hampson, S. E., Andrews, J.A., Barckley, M., Lee, M.E. & Lichtenstein, E. (2003). Assessing perception of synergistic health risk: a comparison of two scales. *Risk Analysis, 23*, 1021-1029.

Harley, N. H., Chittaporn, P., Heikkinen, M.S.A., Meyers, O.A., & Robbins, E.S. (2008). Radon carcinogenesis: Risk data and cellular hits. *Radiation Protection Dosimetry, 130*(1), 107-109.

Hora, S. C., Dodd, N. G., & Hora, J. A. (1993). The use of decomposition in probability assessments of continuous variables. *Journal of Behavioral Decision Making, 6*(2), 133-147. doi: 10.1002/bdm.3960060205

Huppert, H. E., & Sparks, R. S. J. (2006). Extreme natural hazards: population growth, globalization and environmental change. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 364*(1845), 1875-1888. doi: 10.1098/rsta.2006.1803

Institute for Alcohol Studies (2010). *Alcohol and Accidents*: Institute for Alcohol Studies (UK). <http://www.ias.org.uk/resources/factsheets/accidents.pdf> Accessed December 16th, 2011.

Kraus, N., Malmfors, T., & Slovic, P. (1992). Intuitive toxicology: Expert and lay judgments of chemical risks. *Risk Analysis, 12*(2), 215-232.

Laver-Fawcett, A. (2007). *Principles of Assessment and Outcome Measurement for Occupational Therapists and Physiotherapists*. West Sussex, UK.: John Wiley & Sons Ltd.

Levine, T., & Banas, J. (2002). One-Tailed F-tests in communication research. *Communication Monographs, 69*(2), 132-143.

Lin, B. B., Perfecto, I., & Vandermeer, J. (2008). Synergies between Agricultural Intensification and Climate Change Could Create Surprising Vulnerabilities for Crops. *BioScience, 58*(9), 847-854. doi: 10.1641/b580911

Lipkus, I. M. (2007). Numerical, verbal, and visual formats of conveying health risks: suggested best practices and future recommendations. *Medical Decision Making, 27*, 696-713.

Litwin, M.S. (2003). *How to Assess and Intepret Survey Psychometrics* (2nd ed. Vol. 8). California: SAGE Publications.

Office for National Statistics (2009). *Statistics on Alcohol: England, 2009*: Office for National Statistics. <http://www.statistics.gov.uk/cci/nugget.asp?id=1208> Accessed May 4th, 2011.

Okan, Y., Garcia-Retamero, R., Cokely, E. T., & Maldonado, A. (2011). Individual Differences in Graph Literacy: Overcoming Denominator Neglect in Risk Comprehension. *Journal of Behavioral Decision Making*, n/a-n/a. doi: 10.1002/bdm.751

Osterlind, S. J. (1998). *Constructing Test Items: Multiple-Choice, Constructed-Response, Performance, and Other Formats* (2nd ed.). Boston: Kluwer Academic.

Penrod, J., Preston, D.B., Cain, R.E., & Starks, M.T. (2003). A discussion of chain referral as a method of sampling hard-to-reach populations. *Journal of Transcultural Nursing, 14*(2), 100-107.

Pershagen, G., Akerblom, G., Axelson, O., Clavensjo, B., Damber, L., Desai, G., Enflo, A., Lagarde, F., Mellander, H., Svartengren, M., & Swedjemark, G.A. (1994). Residential radon exposure and lung cancer in sweden. *New England Journal of Medicine, 330*, 159-164.

Reips, U.-D. (2000). The web experiment method: Advantages, disadvantages and solutions. In M. H. Birnbaum (Ed.), *Psychological Experiments on the Internet* (pp. 89-118). San Diego, CA: Academic Press.

Rowe, G. W., G. (2001). Differences in expert and lay judgment of risk: myth or reality? *Risk Analysis, 21*, 341-356.

Satterthwaite, D. (2009). The implications of population growth and urbanization for climate change. *Environment and Urbanization, 21*(2), 545-567. doi: 10.1177/0956247809344361

Sjoberg, L. (2000). The methodology of risk perception research. *Quality & Quantity, 34*, 407-418.

Travis, J. M. J. (2003). Climate change and habitat destruction: a deadly anthropogenic cocktail. *Proceeding of the Royal Society of London: Biological Sciences, 270*, 467-473.

Wallsten, T. S., Budescu, D.V., & Zwick, R. (1993). Comparing the calibration and coherence of numerical and verbal probability judgments. *Management Science, 39*(2), 176-190.

Windschitl, P. D. (2002). Judging the accuracy of a likelihood judgment: The case of smoking risk. *Journal of Behavioral Decision Making, 15*(1), 19-35.

Windschitl, P. D. W., G.L. (1996). Measuring psychological uncertainty: Verbal versus numeric methods. *Journal of Experimental Psychology: Applied, 2*(4), 343-364.

Wright, G., Rowe, G., Bolger, F. & Gammack, G. (1994). Coherence, calibration and expertise

in judgmental probability forecasting. *Organisational Behaviour and Human Decision Processes, 57*, 1-25.

Wright, G., Bolger, F. & Rowe, G. (2009). Expert judgment of probability and risk. In K. Samset, Williams, T., & Sunnevag, K.J. (Ed.), *Making Essential Choices with Scant Information: Front-End Decision-Making in Major Projects*. Basingstoke, UK: Palgrave Macmillan.

Wright, G., Saunders, C., & Ayton, P. (1988). The Consistency, Coherence and Calibration of Holistic, Decomposed and Recomposed Judgemental Probability Forecasts. [Article]. *Journal of Forecasting, 7*(3), 185-199.

York, R., Rosa, E. A., & Dietz, T. (2003). STIRPAT, IPAT and ImPACT: analytic tools for unpacking the driving forces of environmental impacts. *Ecological Economics, 46*(3), 351-365. doi: 10.1016/s0921-8009(03)00188-5

Yusuf, H. R., Giles, W.H., Croft, J.B., Anda, R.F., & Casper, M.L. (1998). Impact of multiple risk factor profiles on determining cardiovascular disease risk. *Preventative Medicine, 27*, 1-9.

Footnotes

1. We could not identify a study that autonomously confirmed/denied that the likelihood of having a fatal accident for someone who drives a vehicle whilst intoxicated by alcohol was greater than the sum of the likelihood of a fatal accident for someone who does either of these two activities alone. However, we considered the risk attributable to alcohol-driving to be synergistic. This is because evidence from the Institute for Alcohol Studies (2010) indicates that the annual risk of a fatal accident for a person who becomes intoxicated by alcohol on an average number of occasions is 1 in 100,000. In addition, evidence from the Office for National Statistics (2009) shows that the annual risk of a fatal accident for a person who drives a vehicle on a road, an average number of occasions, is 3 in 100,000. Moreover, Cherpitel et al. (1995) report that the risk of a fatal accident increases 11-fold for a person who drives whilst intoxicated, thus, making the risk of a fatal accident for someone who drives a vehicle whilst intoxicated to be approximately 33 in 100,000; this is greater than the sum of the risk attributable to each constituent hazard in isolation.
2. At the end of the questionnaire participants were asked if they had referred to any external informational sources to help them answer the questionnaire. Two participants stated that they had referred to external sources, so their data was not included in the analysis.
3. Over 82 percent of the participants in Study 1 were either US (n = 45) or UK (n = 42) residents. To assess whether this heterogeneity in residency caused any bias in our data (e.g., only the residents in one of the two countries may have been exposed to a public health campaign that had made them aware of the synergistic risk attributable to one of the three combinations) we examined whether the judgements differed between the participants residing in the UK and those residing in the US. The analysis found no significant difference

between these two groups in their judgments for the alcohol-driving, aspirin-clopidogrel or radon-tobacco combinations (χ^2s (1) <= 1.22, *p*s => 0.35), and there was no significant difference when all three factor combinations were treated as a single variable (χ^2 (1) = 0.48, $p = 0.51$). Hence, we conclude that it is unlikely that this sample heterogeneity led to a bias in our findings.

Table

Table 1. Relative frequencies, as described in each 'numeric frame' judgement task, of adverse outcomes following exposure to/use of a specific risk factor (frequencies calculated from data source shown).

Risk Factor Combination in Judgement Task	***Constituent Risk Factor***	***Adverse Outcome***	***Relative Frequency of Adverse Outcome***	***Data Source(s)***
Alcohol – Driving	*Alcohol*: becoming intoxicated by alcohol an average number of times per year	Fatal accident in any given year	1 in 100,000	Institute for Alcohol Studies (2010)
	Driving: driving vehicle on a road an average number of times per year	Fatal accident in any given year	3 in 100,000	Office for National Statistics (2009)
Aspirin – Clopidogrel	*Aspirin*: taking a low dose each day	Gastro-intestinal bleeding in any given year	100 in 100,000	Hallas et al. (2006)
	Clopidogrel: taking a low does each day	Gastro-intestinal bleeding in any given year	10 in 100,000	Hallas et al. (2006)
Radon – Tobacco	*Radon*: living in a dwelling where there is a high level of radon	Lung cancer during lifetime	4,000 in 100,000	Reif & Heeren (1999)
	Tobacco: smoking 20 cigarettes per day throughout adulthood	Lung cancer during lifetime	17,000 in 100,000	Villeneuve & Mao (1994)

Figures

Figure 1. Study 1: Judgements of the risk model attributable to each risk factor combination (*N* = 106)

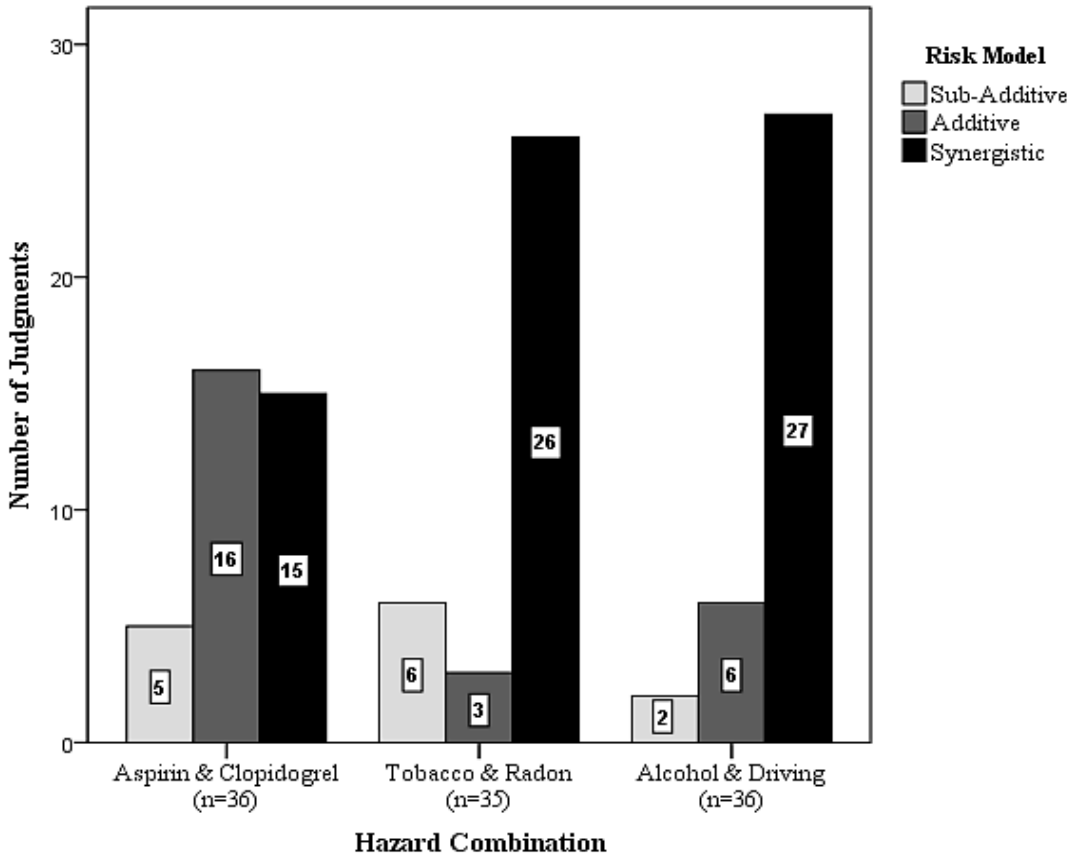
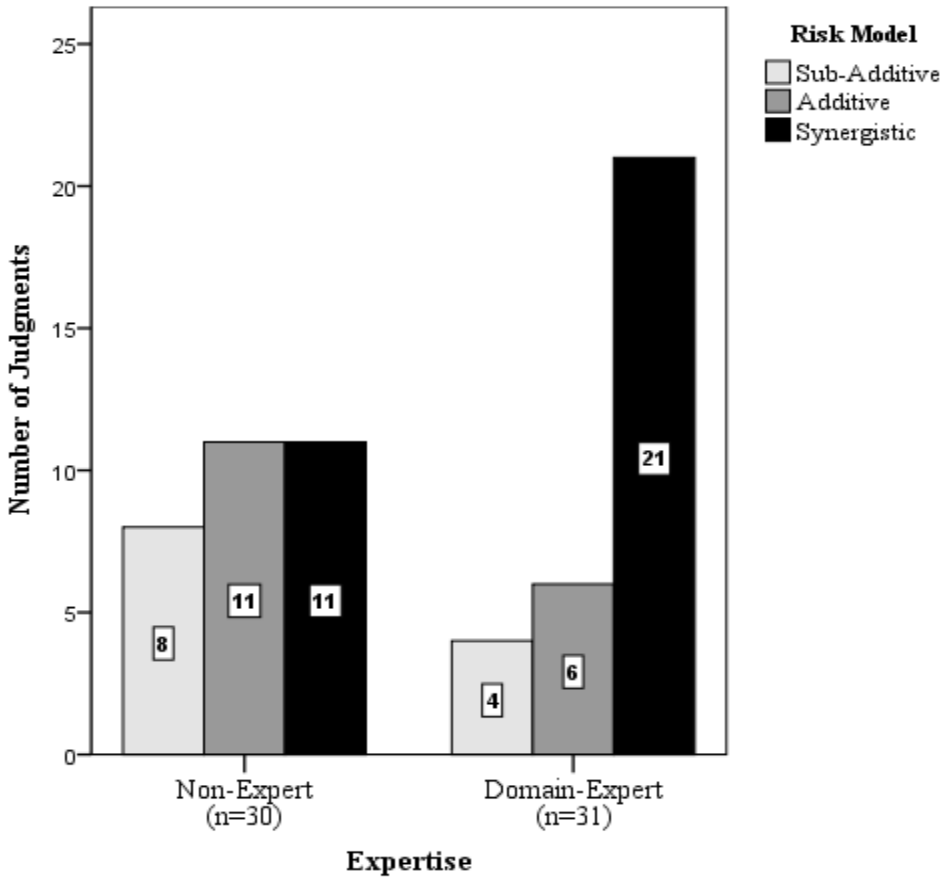


Figure 2. Study 2: Domain-experts' (IPs) and non-experts' judgements of the risk model attributable to the aspirin-clopidogrel drug combination (N = 61).



Appendix

Example judgement task featuring the aspirin-clopidogrel combination in a numeric frame:

Read the paragraph below and respond to the task by ticking one of the boxes.

Research evidence shows that a person who takes a low-dose of 'aspirin' each day has a 100 in 100,000 chance of suffering gastro-intestinal bleeding in any given year. Research evidence also shows that a person who takes a low-dose of the antiplatelet drug 'clopidogrel' each day has a 10 in 100,000 chance of suffering gastro-intestinal bleeding in any given year.

Judgment task: Please now consider the chance of a person suffering gastro-intestinal bleeding in any given year if they take a low-dose of aspirin each day **and** a low-dose of clopidogrel each day.

Do you judge the chance as being either **less than**, **equal to**, or **more than** 'the chance of gastro-intestinal bleeding for a person who takes low-dose aspirin each day' **added to** 'the chance of gastro-intestinal bleeding for a person who takes low-dose clopidogrel each day'?

Less than ☐

Equal to ☐

More than ☐