Conflict and Ambiguity: Preliminary Models and Empirical Tests

Michael Smithson
The Australian National University, Canberra, Australia
Michael.Smithson@anu.edu.au

Abstract

The proposition that conflict and ambiguity are distinct kinds of uncertainty remains debatable, although there is substantial behavioral and some neurological evidence favoring this claim. Recently formal decisional models that combine ambiguity and conflict have been proposed. This paper presents empirical tests of four hypotheses and five models of uncertainty judgments under ambiguity and conflict, via comparisons between pairs of conflicting and ambiguous interval estimates by a sample of 395 adults. The main findings are as follows.

1. Human judges see conflict even in nested intervals with identical midpoints and symmetrically differing endpoints.
2. Identical envelopes of intervals may not be perceived as equally conflictive. Moreover, sets of intervals whose average widths are identical may not be perceived as equally ambiguous.
3. Perceived degree of conflict does not necessarily covary with the magnitudes of the differences between corresponding pairs of interval endpoints. Indeed, a nested pair of intervals may be regarded as more conflictive than a non-nested overlapping pair whose pairs of endpoints differ identically to the nested pair.
4. Judgments of degrees of conflict and ambiguity both contribute independently to judgments of overall uncertainty. However, judgments of ambiguity and conflict appear to be positively correlated.

None of the models pass all empirical tests, but specific suggestions for improving the models are derived from the findings.

Keywords. Uncertainty, ambiguity, conflict, judgment, decision.

1 Introduction

Whether conflict and ambiguity are distinct kinds of uncertainty remains an open question, as does their joint impact on judgments of overall uncertainty. There is behavioral evidence (Smithson 1999, Cabantous 2007, Cabantous et al. 2011, Baillon et al. 2012) and some neurological evidence (Pushskarskaya et al. 2013) in favor of the notion that conflict and ambiguity are separate. However, there are generalized probability frameworks that deal in sets of probabilities, where this distinction appears unnecessary or irrelevant.

Recently formal models of decision making under conflict and ambiguity have been proposed (Gajdos & Vergnaud 2012) that include separate parameters to represent orientations towards conflictive and ambiguous uncertainties. Such models can differ in important ways that are amenable to empirical tests by human judges. Here, we shall examine simple comparisons between interval estimates, where the intervals may or may not overlap, and we will focus on four questions:

1. Do nested intervals (special case: identical midpoints) imply no conflict?
2. Do identical envelopes of intervals imply equal conflict and/or equal ambiguity? What about identical interval averages?
3. Does conflict covary with the magnitudes of the differences between corresponding pairs of interval endpoints?
4. Do judgments of degrees of conflict and ambiguity both contribute independently to judgments of overall uncertainty?

The rationale for questions 1-3 is that conventional pooling rules for sets of quantitative estimates may yield “yes” and “no” answers to these questions. For example, two equally credible interval estimates $[1, 7]$ and $[3, 5]$ may be averaged to yield a pooled interval estimate $[2, 6]$, the same result if both interval estimates were identical intervals $[2, 6]$. So this example could be interpreted as answering "yes" to questions 1 and the average interval version of 2.

A second example, two interval estimates $[1, 5]$ and $[3, 7]$, also may be averaged to yield $[2, 6]$. This example would seem to answer "yes" to question 3 when we compare it to the first example, because in both examples the magnitude of the difference between the lower endpoints is 2 and so is the difference between the upper endpoints. The same comparison also answers "yes" to the identical envelopes version of question 2. But now consider the pair of intervals $[0, 4]$ and $[4, 8]$. Averaging them yields $[2, 6]$ again, despite the fact that their lower and upper endpoints differ by 4 instead of 2. Given that both intervals have the same widths as those in the second example so they are equally ambiguous, it would seem that the degree of conflict does not covary with these differences and this example says "no" to question 3.

A more risk-averse pooling rule that stipulates taking the minimum of the lower endpoints and the maximum of the upper endpoints of equally credible interval estimates says "no" to questions 1 and 2. Pooling intervals $[2, 6]$ and $[3, 5]$ with this rule yields $[2, 6]$, the same result if both interval estimates were identical intervals $[2, 6]$. Clearly the first pair of intervals is, on average, less ambiguous than the second, so perhaps the first pair has some degree of conflict whereas the second identical pair, of course, does not. The lesser ambiguity is then compensated by the greater conflict to yield the same overall uncertainty in the pooled interval. So we have "no" to questions 1 and the identical envelopes version of question 2.

The rationale for question 4 stems from behavioral evidence (Smithson 1999, Cabantous 2007 and Baillon et al. 2012) and recent neurological evidence (Pushkarskaya et al. 2013) that people treat uncertainty arising from conflicting information as distinct from uncertainty arising from ambiguity. Even granting this claim, it is not clear how people combine the two kinds of uncertainty if asked to evaluate the overall uncertainty of a prospect.

Simple empirical tests of all four questions can be constructed by two-alternative forced-choice experiments in conjunction with simple models incorporating each hypothesis. In the next section we shall see that reasonable models of ambiguity and conflict can be constructed to yield "yes" and "no" answers to questions 1-3.

2 Models

Suppose that K judges provide estimates of a quantity of the form $[p_{k1}, p_{k2}, \ldots, p_{kJ}]$, where the p_{kj} are order statistics: $p_{k1} < p_{k2} < \ldots < p_{kJ}$. The simplest setup of this kind, which we shall consider, has two judges, each of whom provides a lower and upper estimate, so that $K = 2$ and $J = 2$.

The k^{th} judge's assessment is ambiguous or vague insofar as the p_{kj} diverge in some sense from one another, and we will consider functions $A(p_{kj})$ to measure ambiguity. Likewise, judges' assessments may conflict with one another insofar as their assessments differ in some sense from each other, and we will also consider functions $C(p_{kj})$ to measure conflict. Finally, a decision maker (DM) who is given these judges' assessments may have a subjective appraisal of the combined uncertainty resulting from both ambiguity and conflict that weighs these two uncertainty components according to their relative aversiveness to the DM. We will therefore investigate uncertainty functions $S(\alpha, \theta, C(p_{kj}), A(p_{kj}))$ that are monotonically increasing in $C(p_{kj})$ and $A(p_{kj})$, where α is the conflict weight and θ is the ambiguity weight.

2.1 Model Types

2.1.1 Variance Component Models

A natural uncertainty metric for both ambiguity and conflict could be variance. Ambiguity effects on judgments and decisions have been explained in terms of variance (Rode et al. 1999), and conflict also has implications for variability in outcomes. The ambiguity of each judge's estimates can be measured by

$$A_k = \sum_{j=1}^{J} (p_{kj} - \overline{p_{k.}})^2 \Big/ J, \qquad (1)$$

so that the total ambiguity is just the within-judge component of the variance of the p_{kj}:

$$A = \sum_{k=1}^{K} A_k / K.$$

An intuitively plausible candidate for measuring conflict, then, is the between-judge variance component:

$$C_1 = \sum_{k=1}^{K} (\overline{p_{k.}} - \overline{p_{..}})^2 \Big/ K \qquad (2)$$

However, an alternative conflict measure is the variance among the order-statistics of the same rank:

$$C_2 = \sum_{k=1}^{K} \sum_{j=1}^{J} (p_{kj} - \overline{p_{.j}})^2 \Big/ JK. \qquad (3)$$

I shall refer to the first model as variance component model 1 (VC1) and the second as VC2. The conflict function in equation 3 differs from that in equation 2 in an important way, because when $\overline{p_{k.}}$ are identical for all K judges, $C_1 = 0$ whereas this is not true for C_2. Thus, VC1 predicts that a pair of interval estimates with identical midpoints will not be perceived as conflictive, whereas VC2 predicts that they will be.

A DM's degree of concern or disutility about ambiguity is represented by a weight, θ, that takes values in the closed unit interval. Likewise, the DM's degree of concern about disagreement or conflict is represented by a weight, α, whose domain also is the unit interval. There are several ways these weights may be employed to combine the ambiguity and conflict measures to construct a measure of overall uncertainty. The simplest is a weighted sum:

$$S(\alpha, \theta, A, C_j) = \theta A + \alpha C_j, \tag{4}$$

where $j = 1, 2$.

2.1.2 Distance Models

Distance models are related to variance models and provide another potential metric for both ambiguity and conflict. A distance model evaluates ambiguity and conflict in terms of distances between order statistics. The ambiguity of the k^{th} judge can be expressed as

$$A_k = \sum_{j_1=1}^{J} \sum_{j_2=1}^{J} |p_{kj_1} - p_{kj_2}|^n / J^2 \tag{5}$$

where $n > 0$ ($n = 2$ is the Euclidean special case). As before, the total ambiguity then is simply

$$A = \sum_{k=1}^{K} A_k / K.$$

Conflict between the judges may be evaluated in two ways. First, we may sum those differences over the ranks and take the absolute value of that sum:

$$C_1 = 2 \sum_{k_1=1}^{K} \sum_{k_2=k_1+1}^{K} \left| \sum_{j=1}^{J} (p_{k_1 j} - p_{k_2 j})^n \right| \Bigg/ K(K-1). \tag{6}$$

Second, we may sum the absolute differences between pairs of order-statistics of the same rank:

$$C_2 = 2 \sum_{k_1=1}^{K} \sum_{k_2=k_1+1}^{K} \sum_{j=1}^{J} |p_{k_1 j} - p_{k_2 j}|^n \Bigg/ K(K-1). \tag{7}$$

I shall refer to the first model as distance model 1 (D1) and the second as D2. As with the previous pair of models, D1 predicts that a pair of interval estimates with identical midpoints will not be perceived as conflictive, whereas D2 predicts that they will be.

As with the variance models, it is possible to combine A and C_j in a weighted sum to produce an overall evaluation of total uncertainty. The result is equation (4) with the weights expressing degrees of disutility regarding ambguity and conflict, and the distance model versions of A and C_j substituted for the variance models versions.

2.1.3 The Gajdos-Vergnaud Model

Gajdos and Vergnaud (2012) develop a model of decision making under ambiguity and conflict based on the Schmeidler-Gilboa (1989) maxmin framework. For the sake of simplicity, I present only the two-state, two-judge special case of their model, and modify their notation to be compatible with the notation used for the other models in this paper. They intended their model to apply to probability judgments; here I extend it to judgments of magnitudes.

In the Gajdos-Vergnaud (GV) model, the α and θ weights are used to modify the order statistics of each judge. The θ parameter contracts the $[p_{k1}, p_{k2}]$ interval around its midpoint at a rate $1 - \theta$, yielding lower and upper bounds

$$\begin{aligned} \pi_{k1} &= p_{k1}(1+\theta)/2 + p_{k2}(1-\theta)/2, \\ \pi_{k2} &= p_{k1}(1-\theta)/2 + p_{k2}(1+\theta)/2. \end{aligned} \tag{8}$$

Gajdos and Vergnaud do not define an ambiguity measure along the lines of those in this paper, but as with the variance and distance models we may construct one by summing the differences $\pi_{k2} - \pi_{k1}$. It can be shown that this ambiguity measure is identical to the distance model's ambiguity measure divided by 2 when $n = 1$.

The GV model treats α as contracting the pairs of interval endpoints p_{kj} and p_{mj} around their mean at the rate $1 - \alpha$. Thus, the order statistics are modified in the following way:

$$\begin{aligned} \gamma_{kj} &= p_{kj}(1+\alpha)/2 + p_{mj}(1-\alpha)/2, \\ \gamma_{mj} &= p_{mj}(1+\alpha)/2 + p_{kj}(1-\alpha)/2. \end{aligned} \tag{9}$$

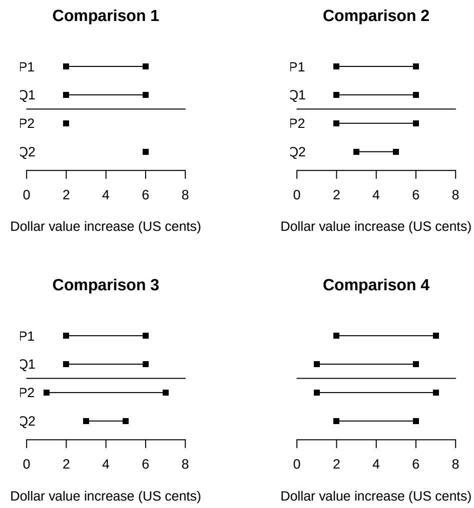
Again, Gajdos and Vergnaud do not define a conflict measure but one may be defined by summing the absolute values of the differences $\gamma_{kj} - \pi_{mj}$. It can be shown that this conflict measure is identical to the distance model's C_2 measure divided by 2 when $n = 1$.

If we evaluate overall uncertainty by summing the ambiguity and conflict measures, clearly we obtain an uncertainty measure identical to that in D2 when $n = 1$.

An alternative evaluation of overall uncertainty is suggested by the maxmin decisional model incorporated into the GV framework. In the development of the GV decisional model, the order statistics are transformed by one parameter and then those results are transformed in turn by the second parameter, according to equations (8) and (9). It is not difficult to show that this procedure is commutative, so that if the α transformation occurs before or after the θ transformation the result is the same. Thus, we may define our alternative uncertainty measure by

$$S(\alpha, \theta, A, C) = \max_{k,j}(\gamma_{kj}) - \min_{k,j}(\gamma_{kj}). \qquad (10)$$

As will be demonstrated, this measure does not behave identically to the measure for D2.



3 Method

Hypotheses and the models were tested via an online experiment. The online study was reviewed and approved by the Australian National University Human Research Ethics Committee. The participant sample consisted of 508 North American adults (205 women, 189 men, 1 unspecified; with mean age = 39.95, sd = 15.04), recruited through Qualtrics, of which 395 cases were found to be trustworthy data. Four comparisons between two pairs of estimates, $\{P_1, Q_1\}$ and $\{P_2, Q_2\}$, were used to test questions 1-3, their results also lending insight into question 4. Comparisons 2 and 3 test question 1, Comparisons 3 and 4 test question 2, and Comparisons 2-4 partially test question 3. These comparisons are graphed in Figure 1. Participants were presented with both the graphs and verbal statements of the estimate pairs. They were asked to choose which pair of estimates exhibited more agreement, which exhibited more ambiguity, and which made them feel more uncertain about the quantity being estimated.

There were two conditions, differing solely on the nature of the estimate. In one condition they were told that the estimates were experts' predictions of the change in global average temperature by the year 2040 (in degrees Celsius). In the other, they were told the estimates were experts' predictions of the change in the value of the Australian dollar against the American dollar in the next 5 years (in US cents). Both scenarios are fictitious, and participants were advised of this in a debriefing at the end of the online survey. Neither of these scenarios is based on expert predictions. The global average temperature scenarios actually are over-estimates of warming, according to the IPCC (2007) report, and the estimates reported therein do not disagree as much as the estimates in some of these scenarios do. Genuine forecasts of currency fluctuations seldom range farther into the future than 6 months to one year, and near-term predictions for the Australian dollar's exchange-rate against the US dollar are mixed with some predicting a decline and others predicting an increase. The goal here was to provide identical numbers under the guise of very different topics, to ascertain whether topics might influence perceptions of conflict or ambiguity. All of this having been said, the topic of the estimates turned out to make no significant difference to people's choices, so from here on these two conditions are ignored.

An example of the text of the first condition is presented here.

> In this section, we want you to make some judgments about estimates of the increase in average temperature by the year 2040. You will be presented with two pairs of estimates from refereed climate science forecasts. We are interested in which pair you think has the greatest uncertainty.
> Expert P1: By 2040 global average temperature will have increased by 2-6 degrees Celsius
> Expert Q1: By 2040 global average temperature will have increased by 2-6 degrees Celsius
> Expert P2: By 2040 global average temperature will have increased by 2 degrees Celsius
> Expert Q2: By 2040 global average temperature will have increased by 6 degrees Celsius
> Taken together, which pair of experts do you think is in more agreement?
> Taken together, which pair of experts do you think is more vague?

Taken together, which pair of experts makes you more uncertain about the temperature increase?

The models presented in the previous section all agree that in Comparison 1, $\{P_1, Q_1\}$ is more ambiguous and less conflictive than $\{P_2, Q_2\}$, and whether one is rated as more uncertain overall depends on the magnitudes of the α and θ parameters. For Comparison 2, all models agree that $\{P_1, Q_1\}$ is more ambiguous than $\{P_2, Q_2\}$, but GV, D2, and VC2 rate $\{P_1, Q_1\}$ as less conflictive than $\{P_2, Q_2\}$ whereas D1 and VC1 rate them as equally conflictive. In Comparison 3 the models make the same predictions about conflict as in Comparison 2, but while VC1 and VC2 rate $\{P_1, Q_1\}$ as less ambiguous than $\{P_2, Q_2\}$, GV, D1 and D2 rate them as equally ambiguous. Finally, for Comparison 4, the models' predictions regarding ambiguity are the same as in Comparison 3, but D1 and VC1 rate $\{P_1, Q_1\}$ as more conflictive than $\{P_2, Q_2\}$ whereas GV, D2, and VC2 rate them as equally conflictive.

Overall uncertainty predictions from the models are not determined for all four comparisons because they may vary with the α and θ parameters. Nevertheless, for every model at least two comparisons yield fixed outcomes. In Comparison 2, GV, D1 and VC1 rate $\{P_1, Q_1\}$ as more uncertain than $\{P_2, Q_2\}$. In Comparison 3, all models except D1 rate $\{P_1, Q_1\}$ as less uncertain than $\{P_2, Q_2\}$; D1 rates them as equally uncertain. In Comparison 4, GV amd D1 rate $\{P_1, Q_1\}$ as more uncertain than $\{P_2, Q_2\}$, VC1 amd DVC2 rate $\{P_1, Q_1\}$ as less uncertain than $\{P_2, Q_2\}$, and D2 rates them as equally uncertain.

4 Results

4.1 Questions 1-3

Regarding question 1, in Comparisons 2 and 3 large majorities of respondents chose the nested interval pair as being more conflictive than the identical interval pair. For Comparison 2, 83.8% made this choice (95% confidence interval (CI) = $[79.8\%, 87.1\%]$); and for Comparison 3, 87.6% made this choice (95% CI = $[84.0\%, 90.5\%]$). These figures are similar to the percentage choosing the two pointwise estimates in Comparison 1 as more conflictive than the identical intervals (84.3%). An unexpected finding was that in Comparison 4, 61.5% chose the nested interval pair as more conflictive than the non-nested, overlapping pair (95% CI = $[56.6\%, 66.2\%]$). These results all strongly suggest that nested interval estimates are perceived as conflictive even when they have identical midpoints.

The finding regarding conflict in Comparison 4 also addresses questions 2 and 3, indicating that neither identical envelopes nor equal differences between pairs of endpoints will ensure that pairs of estimates will be regarded as equally conflicting. The finding for Comparison 3 demonstrates that identical average interval widths for pairs of estimates also will not ensure that they are perceived as equally conflicting.

Question 2 applied to ambiguity, on the other hand, yielded mixed results. In Comparison 4, where the pairs have identical envelopes, we cannot rule out the possibility that respondents were evenly split on which pair is the more ambiguous (95% CI = $[47.0\%, 56.8\%]$). However, in Comparison 3 where the average interval widths are the same for both pairs, 78.0% chose the nonidentical pair of intervals as more ambiguous than the identical pair (95% CI = $[73.6\%, 81.8\%]$).

Question 3 also can be addressed via a test for marginal homogeneity in the cross-classification of Comparison 2 and 3 choices regarding conflict. In Comparison 2 the first pair of intervals was chosen as more agreeing by 83.8% of respondents and in Comparison 3 87.6% chose the first pair. Because the widths of the nested pair in Comparison 2 differ by less than those in Comparison 3 we should indeed expect a higher percentage in Comparison 3. However, a 95% CI for the paired difference yields $[-0.11\%, 7.75\%]$ so we fail to reject the null hypothesis of no difference.

4.2 Overall Uncertainty and Question 4

Comparison 1 offers indirect corroboration of Smithson's (1999) conflict aversion hypothesis, because 60.0% of respondents chose the pointwise pair of disagreeing estimates as more uncertain than the pair of agreeing interval estimates (95% CI = $[54.1\%, 63.7\%]$). A similar percentage, 58.0%, chose the nested pair of intervals in Comparison 2 as more uncertain than the pair of agreeing interval estimates (95% CI = $[53.1\%, 62.7\%]$), and a substantially greater percentage, 76.7%, made the same choice in Comparison 3 (95% CI = $[72.3\%, 80.6\%]$). Finally, in Comparison 4 55.2% chose the nested pair of intervals as more uncertain than the overlapping pair (95% CI = $[50.3\%, 60.0\%]$). These latter three findings indicate that both perceived conflict and ambiguity may be independently contributing to overall perceived uncertainty.

A direct test of this (i.e., question 4) is a mixed logistic regression model utilizing the data from all four comparisons. This model included main-effects terms for comparisons, ambiguity and agreement (conflict), with random effects for the latter two covariates. A model with interaction terms did not improve fit significantly ($\chi^2(6) = 9.642, p = .141$). Both the agreement and ambiguity terms were significant in the expected directions ($z = -6.576, p < .0005$ and $z = 12.568, p < .0005$, respectively), the ambiguity effect being nearly twice as large.

4.3 Model Performance

The performance of the five models can be evaluated in two ways. First, we can simply assign each a "pass" or "fail" grade for every prediction made by each model regarding comparative conflict, ambiguity, or uncertainty. Second, for ambiguity and conflict we may use the differences between the scores each model assigns to every relevant pair of estimates to predict respondent choices via mixed logistic regressions.

Table 1 summarizes the "pass" or "fail" results. Only Comparisons 2-4 are shown because all models passed Comparison 1 on conflict and ambiguity and made no determinate predictions for uncertainty. "P" indicates that the model's prediction is in accordance with the empirical result; "F" indicates that the model's prediction is the opposite of the result; "N" that the model's prediction is equality whereas the result suggests a difference; and "U" that the status of the model's prediction is undetermined by the result because the null hypothesis could not be rejected.

Table 1: Model Pass-Fail Results

	Conflict			Ambiguity			Uncertainty		
	2	3	4	2	3	4	2	3	4
GV	P	P	N	U	N	P	F	P	F
D1	N	N	F	U	N	P	F	N	F
D2	P	P	N	U	N	P		P	N
VC1	N	N	F	U	P	U	F	P	F
VC2	P	P	N	U	P	U		P	P

Beginning with the conflict results, models D1 and VC1 fail three of the four comparisons because they are the models predicting that pairs of intervals with identical midpoints will not be considered to be conflicting. The other models pass three of the four comparisons. None of the models pass Comparison 4. The ambiguity results are equivocal, with all models passing two comparisons and none performing markedly better than the others. The uncertainty results also are mixed. No model with a determinate prediction passes Comparison 2, all but one pass Comparison 3, and only one (VC2) passes Comparison 4.

We now turn to the mixed logistic regressions. The D1 and VC1 models' conflict scores for the five distinct pairs of estimates used in the comparisons are proportional to one another, and the GV, D2 and VC2 models' conflict scores are proportional to one another. So there are two mixed logistic regressions to compare: The D1-VC1 and GV-D2-VC2 models. The log-likelihood of the GV- D2-VC2 model is -1073.39 and, as might be expected, markedly higher than the log-likelihood of the D1-VC1 model (-1083.40).

As with the conflict scores, the ambiguity scores for the D1 and VC1 models are proportional to one another and the ambiguity scores for the GV, D2 and VC2 models are proportional to one another. The The log-likelihoods of the GV- D2-VC2 and D1-VC1 models are fairly similar (-1336.76 and -1332.28 respectively).

5 Discussion

The results strongly indicate that the answer to question 1 is "no", at least for the rather de-contextualized comparisons used in this study. Even so, I urge caution regarding generalizability, having witnessed at least one applied context (a consultancy with a banking organization) in which stakeholders decided that nested estimates should *not* be considered as disagreeing. Other contextual factors could alter the answer to this question. Fior example, it is plausible that if one estimate is known to be based on a larger data set than the other, nested intervals might not be taken to indicate disagreement but instead attributed to the different sizes of the data sets.

Likewise, the conflict comparison results suggest the answer to question 2 is "no". However, the ambiguity comparisons are inconclusive regarding this question, and further investigations will be required to ascertain the conditions under which identical envelopes of intervals confer equal ambiguity. As indicated above, additional information about the basis for the estimates could alter this outcome as well.

Question 3 also has been answered in the negative, both in the failure to find a significant difference between choices in Comparisons 2 and 3, and in another unexpected fashion. None of the models or pooling rule considerations anticipated the finding in Comparison 4 that a nested pair of intervals would be regarded as more conflictive than a non-nested overlapping pair whose pairs of endpoints differed identically to the nested pair. This finding begs for interpretation, and that will be addressed shortly.

The mixed logistic regression demonstrated that both conflict and ambiguity choices made independent contributions to predicting uncertainty choices between pairs of estimates in the four comparisons. Moreover, the type of comparison did not significantly moderate the effect of either conflict or ambiguity. Thus, respondents generally behaved as though they perceived ambiguity and conflict as distinct contributors to overall uncertainty.

Nonetheless, this result suggests another question, namely whether ambiguity and conflict choices are associated. In this sample, judgments of ambiguity and agreement are strongly negatively related for all four comparisons (i.e., ambiguity and conflict are positively associated). That is, the odds of choosing $\{P_1, Q_1\}$ as the more ambiguous pair are higher if the respondent also chose $\{P_2, Q_2\}$ as the more agreeable (and vice versa). The odds-ratios for Comparisons 1, 2, 3, and 4 are 2.90, 6.79, 14.13, and 22.84, respectively. For Comparison 1 this finding is somewhat surprising because the pairs of estimates are constructed so that one pair is clearly ambiguous and the other clearly conflicting. It is not as surprising for Comparisons 2 and 4 because there is no definite majority view on which pair of estimates is the more ambiguous in either comparison. However, it is unsurprising for Comparison 3 because substantial majorities of respondents chose the second pair of estimates, $\{P_2, Q_2\}$, as more ambiguous and the first pair as showing more agreement.

The consistency of this positive relationship suggests that people may regard conflict and ambiguity as entailing one another: The greater the perceived conflict, the greater the perceived ambiguity, and vice-versa. This is not an irrational association to make, given that there are situations where ambiguity can generate conflict or conflict can generate ambiguity.

Table 2 displays the crosstabulations of the choices for all four comparisons. The negative association between the ambiguity and conflict choices is especially clear in Comparisons 2-4, where the majority of respondents who have chosen $\{P_1, Q_1\}$ as the more agreeing pair also have chosen $\{P_2, Q_2\}$ as the more ambiguous, while the majority who have chosen $\{P_2, Q_2\}$ as the more agreeing have chosen $\{P_1, Q_1\}$ as the more ambiguous.

Finally, let us consider the issue of modeling conflict and ambiguity jointly. Starting with ambiguity, as mentioned earlier, none of the models were clearly superior to the others in predicting ambiguity choices. The GV, D2 and VC2 models differ from the D1 and VC1 models in their predictions for Comparisons 3 and 4, so that the first three pass Comparison 4 while the latter two pass Comparison 3. Inspection of Table 2 reveals that D1 and VC1 pass both Comparisons 3 and 4 for those people who chose the first pair of intervals in each comparison as showing more agreement. As mentioned above, in Comparisons 2-4

Table 2: Ambiguity-Agreement Association

	Agreement		
Ambig.	$\{P_1, Q_1\}$	$\{P_2, Q_2\}$	
$\{P_1, Q_1\}$	173	47	
$\{P_2, Q_2\}$	160	15	Comparison 1
	$\{P_1, Q_1\}$	$\{P_2, Q_2\}$	
$\{P_1, Q_1\}$	129	52	
$\{P_2, Q_2\}$	202	12	Comparison 2
	$\{P_1, Q_1\}$	$\{P_2, Q_2\}$	
$\{P_1, Q_1\}$	52	35	
$\{P_2, Q_2\}$	294	14	Comparison 3
	$\{P_1, Q_1\}$	$\{P_2, Q_2\}$	
$\{P_1, Q_1\}$	57	133	
$\{P_2, Q_2\}$	186	19	Comparison 4

the majority choice of which pair is more ambiguous switches depending on which pair is seen as showing more agreement. The clear suggestion is to build and test models of conflict and ambiguity assessment that take this positive relationship into account.

Turning now to conflict, the GV, D2 and VC2 models perform markedly better than the D1 and VC1 models in predicting conflict choices because the latter two models consider nested interval estimates as having no conflict. However, none of the models passed Comparison 4.

One interpretation of the respondents' conflict choices in Comparisons 2-4 is that some people may perceive differences in interval widths as indicating disagreement. Thus, the second pair of estimates in Comparison 4 is doubly penalized for conflict because the endpoints differ and so do the interval widths, whereas in the first pair the endpoints differ by the same amounts but the interval widths agree (i.e., the experts are equally vague).

It is not difficult to amend the conflict models presented thus far to accommodate a penalty for differing vagueness. In the distance and variance component models it simply amounts to adding a distance measure and a variance component, respectively, that accounts for differences in interval widths. Doing so does not alter their predictions for any of the other comparisons, so they now pass Comparison 4. Moreover, their mixed logistic regression log-likelihoods are markedly better than their original counterparts (-1062.63 and -1064.75). The new models also present novel predictions regarding other comparisons and thus suggest specific tests of their validity. These will be undertaken in future experiments.

Readers will have noticed that the estimate scenarios in this study were considerably simplified, omitting any information about how the experts arrived at their estimates, the data on which the estimates were based, the experts' qualifications, and so on. As mentioned earlier in this section, such information can affect perceptions of conflict and ambiguity. For instance, two differing estimates based on separate analyses of the same data set would be likely to be percieved as a more striking conflict than the same two estimates based on separate (but, say, equal-sized) data sets. Likewise, knowledge of two experts' prior (dis)agreements with one another on similar issues could substantially influence perceptions of how strong their current disagreement is. Examples of factors potentially affecting perceptions of ambiguity are the amounts of evidence on which estimates are based and the level of relevant expertise possessed by the estimator. Finally, relevant perceiver characteristics include tolerance of uncertainty, agreeableness, need for closure, and prior alignment with one or another expert's position on issues relevant to the estimates. There is considerable scope, therefore, for experimentally investigating the effects of particular kinds of information and assessing the impacts of psychological covariates on perceptions of conflict and ambiguity.

References

[1] Baillon, A., Cabantous, L. & Wakker, P. (2012) Aggregating imprecise or conflicting beliefs: An experimental investigation using modern ambiguity theories. *Journal of Risk and Uncertainty*, 44, 115-147.

[2] Cabantous, L. (2007) Ambiguity aversion in the field of insurance: insurers attitude to imprecise and conflicting probability estimates. *Theory and Decision*, 62, 219240.

[3] Cabantous, L., Hilton, D., Kunreuther, H., & Michel-Kerjan, E. (2011) Is imprecise knowledge better than conflicting expertise? Evidence from insurers decisions in the United States. *Journal of Risk and Uncertainty*, 42, 211-232.

[4] Gajdos, T. & Vergnaud, J-C. (2012) Decisions with conflicting and imprecise information. *Social Choice and Welfare*, in press.

[5] Intergovernmental Panel on Climate Change, *Summary for policymakers: Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Retrieved May 2010 from http://www.ipcc.ch/pdf/assessment-report/ar4/wg1/ar4-wg1-spm.pdf(2007).

[6] Pushkarskaya, H., Smithson, M., Joseph, J.E., Corbly, C., & Levy, I. (2013) Decision making under ambiguity and conflict. *Unpublished manuscript.*

[7] Rode, C., Cosmides, L., Hell, W., & Tooby, J. (1999). When and why do people avoid unknown probabilities in decisions under uncertainty? Testing some predictions from optimal foraging theory. *Cognition*, 72, 269304.

[8] Smithson, M. (1999). Conflict aversion: preference for ambiguity vs. conflict in sources and evidence. *Organizational Behavior and Human Decision Processes*, 79, 179198.