THINKING & REASONING, 2012, 18 (3), 379–393

Psychology Press
Taylor & Francis Group

Utility conditionals as consequential arguments: A random sampling experiment

Jean-François Bonnefon

CNRS and Université de Toulouse, Toulouse, France

Research on reasoning about consequential arguments has been an active but piecemeal enterprise. Previous research considered in depth some subclasses of consequential arguments, but further understanding of consequential arguments requires that we address their greater variety, avoiding the risk of over-generalisation from specific examples. Ideally we ought to be able to systematically generate the set of consequential arguments, and then engage in random sampling of stimuli within that set. The current article aims at making steps in that direction, using the theory of utility conditionals as a way to generate a large set of consequential arguments, and offering one study illustrating how the theory can be used for the random sampling of stimuli. It is expected that further use of this method will bring more diversity to experimental research on consequential arguments, and more robustness to models of argumentation from consequences.

Keywords: Argument; Conditional; Preference; Sampling; Utility.

Argumentation models have the advantage of simultaneously addressing reasoning and decision making. Arguments can inform conclusions, and be relevant to reasoning (Bonnefon, 2004; Corner & Hahn, 2009; Hahn & Oaksford, 2007; Rahwan, Madakkatel, Bonnefon, Awan, & Abdallah, 2010; Thompson & Evans, 2012 this issue); but they can also inform choices, and be relevant to decision making (Amgoud, Bonnefon, & Prade, 2005; Bonnefon, Dubois, Fargier, & Leblois, 2008; Corner, Hahn, & Oaksford, 2011; Dubois, Fargier, & Bonnefon, 2008).

Correspondence should be addressed to Jean-François Bonnefon, CLLE (CNRS, UTM, EPHE), Maison de la recherche, 5 allées A. Machado, 31058 Toulouse, Cedex 9, France. E-mail: bonnefon@univ-tlse2.fr

http://www.psypress.com/tar http://dx.doi.org/10.1080/13546783.2012.670751

Because they can inform both conclusions and decisions, arguments draw equally on beliefs and preferences. In simple cases an argument warrants a belief on the basis of another belief (1-a), or a preference on the basis of another preference (1-b):

(1) (a) I did well on the test, therefore I'll pass the exam.
(b) She doesn't like French cuisine, therefore she prefers not to go to a French restaurant.

Other combinations are possible, though. Of particular interest in this article are the situations where preferences are used to infer beliefs. Under their degenerated form these arguments lead to a wishful thinking fallacy known as *argumentum ad consequentiam* (Walton, 1999), as in the following example:

(2) If there is no afterlife, then life is meaningless. We do not want life to be meaningless. Therefore, there is an afterlife.

Not all consequential arguments are this fallacious though. Indeed, compare (2) to (3):

(3) If Amy eats oysters, she gets very sick. She does not want to be very sick. Therefore, she does not eat oysters.

Argument (3) sounds perfectly reasonable, despite being very similar in structure to Argument (2). The first sentence alone, "If Amy eats oysters, she gets very sick" seems to be sufficient for us to spontaneously infer that she does not and will not have oysters. Many other examples can be constructed, such as:

(4) (a) If I help you, my boss will fire me.
(b) If you revise the paper, we will publish it.
(c) If Elsie buys the DVD, she will make her children happy.

All three sentences appear to support inferences about what is going to happen: the speaker is not going to help the hearer (4-a), the hearer is going to revise the paper (4-b), Elsie is going to buy the DVD (4-c). What is more, these inferences appear to be persuasive enough to supersede logical conclusions. Consider for example the following problem:

(5) If Zoe is invited to the party, she'll buy a new dress. Zoe is invited to the party.

A typical sample of reasoners would almost unanimously conclude that Zoe will buy a new dress. Indeed, the *Modus Ponens* inference from "if p then q, p" to "q" is probably the single most endorsed inference in reasoning experiments. But now compare (5) to (6):

(6) If Zoe is invited to the party, she'll buy a new dress. Zoe is invited to the party. If she buys a new dress, she can't pay the rent.

Now the conclusion that Zoe will buy a new dress does not seem as solid as before, and most reasoners would in fact reject it (Bonnefon & Hilton, 2004). In sum, arguments can be constructed where the existence of a preference warrants an inference about a state of the world, and these arguments can be more persuasive than even the most persuasive of logical arguments, as evidenced by the rejection of the *Modus Ponens* inference from Example (6).

Research on reasoning about consequential arguments has been an active but piecemeal enterprise. As reviewed in the next section, previous research considered in depth some specific subclasses of consequential arguments, most notably slippery slope arguments, and various types of inducements. While this focused research has been highly fruitful, further understanding of consequential arguments requires that we address their greater variety, avoiding the risk of over-generalisation from specific examples. Ideally we ought to be able to systematically generate the set of all consequential arguments, and then engage in random sampling of stimuli within that set. The current article aims at making a few steps in that direction, using the theory of utility conditionals (Bonnefon, 2009; Bonnefon, Girotto, & Legrenzi, 2012) as a way to generate a large set of consequential arguments and offering one study illustrating how the theory can be used for the random sampling of stimuli.

CHARTED AND UNCHARTED TERRITORIES

This section briefly reviews previous empirical work on consequential arguments, focusing on arguments expressed in conditional form. As we will see, most previous research focused on two subsets of arguments, namely slippery slope arguments and inducive speech acts. We will conclude the section by considering the greater variety of consequential arguments that are left for future research, and the need to rely on a systematic strategy for generating arguments that are not already (and conveniently) lexicalised.

Slippery slopes

Slippery slope arguments are used to reject a proposal based on the belief that it would potentially lead to the re-evaluation of a position that is currently deemed undesirable, for example (adapted from Corner et al., 2011):

(7) (a) If we allow the cloning of farm animals, in the future people will want to clone humans as a form of ethnic cleansing.
(b) If we categorise assault in possession of a knife as a minor offence, we will be led to categorise assault in possession of a gun as a minor offence.

In these two examples, an argument is made against accepting x (cloning farm animals, categorising assault in possession of a knife as a minor offence), because it would increase the probability that an undesirable proposal y be accepted in the future (cloning humans as a form of ethnic cleansing, categorising assault in possession of a gun as a minor offence).

The Bayesian approach to slippery slopes arguments (Corner & Hahn, 2007; Corner et al., 2011; Hahn & Oaksford, 2007) assumes that their perceived persuasiveness is determined by two parameters and one psychological mechanism.[1] The persuasiveness of a slippery slope argument should depend on the disutility of its outcome, and on the conditional probability of this outcome given the acceptance of the proposal x. Furthermore, this probability is claimed to derive from a mechanism of category boundary re-appraisal. In essence, to consider that x belongs to a given category (e.g., acceptable policies, minor offences, tolerable risks) amounts to redefining the border of that category in order to include x, and the more similar x is to y, the more plausible it is that this re-appraisal will move y within the boundaries of that category. Therefore the more similar x is to y, the greater the probability of accepting y given that x is accepted, and the more persuasive the slippery slope argument.

Corner et al. (2011) offered empirical evidence for this Bayesian approach. They showed that the utility of y and the conditional probability of accepting y when accepting x both contributed to the persuasiveness of a slippery slope argument; and that increasing the similarity of x and y increased the perceived persuasiveness of a slippery slope argument. Note how these findings are presumably different in terms of their generalisation to all consequential arguments: the effect of the similarity of x and y is probably specific to slippery slope arguments—whereas the effect of utility and probability might be generalised to a larger class of consequential arguments, as we will now consider by turning our attention to inducive speech acts.

[1]This Bayesian approach is not limited to slippery slope arguments; see Hahn and Oaksford (2006) for other consequentialist arguments.

Inducive speech acts

Most research on consequential arguments focused on inducive speech acts such as promises, threats, tips, and warnings:

(8) (a) If you wash the dishes I'll take out the trash.
(b) If you interrupt me one more time I'll expel you.
(c) If you take the A train you'll be there faster.
(d) If you have another drink you won't be able to drive.

All these examples invite inferences from what we assume to be the preferences of the listener. For example, our confidence in the conclusion that the listener of (8-a) will wash the dishes is apparently a function of how much stronger is her preference for washing the dishes as compared to her preference for taking out the trash, as well as a function of the probability that the speaker will indeed take out the trash if the listener washes the dishes. Likewise, our confidence in the conclusion that the listener of (8-c) will take the A train is a function of how costly it would be to do so, as compared to the benefits of a faster trip, as well as a function of how likely it is that train A is indeed faster (Evans, Neilens, Handley, & Over, 2008; Evans & Twyman-Musgrove, 1998; Hoeken, Timmers, & Schellens, 2012 this issue; Ohm & Thompson, 2004, 2006; Verbrugge, Dieussaert, Schaeken, & Van Belle, 2004).

Another trend of research has tried to define the characteristics that turn a conditional "if p then q" into a promise (8-a), a threat (8-b), a tip (8-c), or a warning (8-d). All these definitions feature broadly the same elements: p is an action of a listener, q is valued by the listener (positively for promises and tips, negatively for threats and warnings), and whether q is an action of the speaker makes the difference between promises and tips, or between threats and warnings (Amgoud, Bonnefon, & Prade, 2007; Evans, 2005; Haigh, Stewart, Wood, & Connell, 2011; López-Rousseau & Ketelaar, 2006). Some accounts include additional conditions, such as p being valued by the speaker, positively for promises, and negatively for threats (Beller, Bender, & Kuhnmünch, 2005; Legrenzi, Politzer, & Girotto, 1996).

This definitional effort is important for our current purpose, because it calls attention to the fact that research on consequential arguments has focused so far on a small subset of these arguments. There are a vast number of "if p, then q" conditionals where p and q are preferred to various degrees by various agents, which do not fall in the category of inducements or slippery slope arguments. Various examples are introduced in the next section, where we eventually call for a reversal of research strategy. Previous research started from already identified consequential arguments, and worked towards defining their structure—but in so doing we have no guarantee that our results are not limited to a tiny subset of all consequential arguments. To allay this concern we need a method to systematically generate consequential arguments of all structures, regardless of whether they already correspond to a known speech act or fallacy.

Uncharted territories

Many conditionals appear to be used as consequential arguments, which cannot be categorised as promises (threats, etc.) or slippery slopes. A few of them were previously identified, for example consequential conditionals (9-a) (Bonnefon & Hilton, 2004), or persuasion conditionals (9-b) (Thompson, Evans, & Handley, 2005):

(9) (a) If Sophie takes this drug, she will be cured;
(b) If the Kyoto accord is ratified, then there will be a severe downturn in the economy.

The consequential conditional (9-a) invites the inference that Sophie will take the drug, and the persuasion conditional (9-b) invites the inference that the speaker believes the Kyoto accord should not be ratified. Just as promises (threats, etc.), "if p then q" consequential and persuasion conditionals were defined in terms of the preferences of various agents over the occurrence of p and q, and the control that various agents have over the realisation of p and q. For consequential conditionals a third party (neither the speaker nor the listener) has control over p, and this same agent has preferences over q. For persuasion conditionals a third party has control over p, and some agents (not specifically but possibly the third party, the speaker or the listener) have preferences over q.

Consequential and persuasion conditionals allow for greater variety in the preference structure of consequential arguments, but they are still very far from covering the full set of potential structures. Consider these three examples:

(10) (a) If I let you go away, my boss will fire me.
(b) If you give me a bonus, she'll sue you in court.
(c) If the customers ask for the rebate I promised, I'll go bankrupt.

All three conditionals appear to invite conclusions based on the preferences of the agents involved: the first speaker is not going to let the listener go away, the second speaker is not going to get a bonus, and the third speaker is presumably about to go bankrupt. All three conditionals thus appear to be consequential arguments, albeit they do not fit any of the categories we have encountered so far. A full perspective on consequential arguments would require investigation of these argument forms, but not only these. Ideally a full perspective on consequential arguments would require a procedure to systematically generate potential argument forms, eliminating the need to generate them by intuitive trial and error.

Such a systematic procedure would allow us to engage in random sampling of stimuli within the set of (conditional) consequential arguments, rather than to focus our investigations on specific subsets. There is nothing wrong, of course, in focusing on specific subsets of consequential arguments when these subsets have intrinsic interest—and this is arguably the case for slippery slope arguments and inducive speech acts. My proposal here is rather to give us the possibility to engage *in parallel* in random sampling of consequential arguments, alongside focused investigations of charted and yet uncharted subsets of interest. One available tool to generate consequential arguments, to which we now turn, is the theory of utility conditionals.

UTILITY CONDITIONALS

Utility conditionals are "if p then q" conditionals such that p or q or both bear on the utilities (preferences) of some agents. Bonnefon (2009) offered a theory of utility conditionals and the inferences people draw from them. The theory consists of a representational tool (the utility grid), which summarises in compact form the utilitarian features of a conditional statement, along with a set of folk axioms of decision that captures reasoners' likely beliefs about the way most agents make their decisions. Applying the folk axioms of decision to the utility grid of a conditional allows prediction of the conclusions it invites by virtue of being a consequential argument. The utility grid of a conditional statement has the following general form:

$$\begin{Bmatrix} x & u & y \\ x' & u' & y' \end{Bmatrix}.$$

The first row of the grid contains the information related to the *if*-clause of the conditional. That is, it displays the agent x (left column) who can potentially take action p, and the utility u (central column) that this action would have for a given agent y (right column). The second row of the grid contains the corresponding information with respect to the *then*-clause of the conditional. That is, it displays the agent x' (left column) who can potentially take action q, and the utility u' (central column) that this action would have for a given agent y' (right column).

The set of all agents is denoted by $\mathcal{A}$. By convention, the agent who states the conditional is denoted by s (for "speaker"), the agent at whom the conditional is directed is denoted by h (for "hearer"), and e (for "someone else") denotes an agent who is neither the speaker nor the hearer. When p or q is not an action that can be taken by an intentional agent but is rather an event or a state of the world, it is noted as being undertaken by a special, neutral agent ω. The agent ω can be thought as "the world" or the body of laws that govern the world.

For the sake of simplicity, utility is represented in the grid by its sign: u and u' take their values from $\{-, 0, +\}$, where $-$ and $+$ respectively stand for any significantly negative and positive values. Note that $u = 0$ means that action p is not known to have any utility for any agent. By convention, such an action has $\mathcal{A}$ as a target ($\mathcal{A}$, as previously noted, is the whole set of agents).[2]

Utility conditionals are consequential arguments, whose conclusions the theory predicts by applying folk axioms of decisions to utility grids. Folk axioms of decision capture the naive understanding that people have of the way other agents make their decisions (Miller, 1999; Smedslund, 1997). The folk axiom of limited altruism, for example, states that agents take actions that increase the utility of others, as long as their own utility does not decrease as a consequence. If people use this folk axiom to predict the decisions of other agents, then their inferences can be predicted by looking for specific configurations of the utility grid. Consider for example this general configuration, where the black dot stands for any value of the corresponding parameter, and where $(u, z) \neq (-, x)$:

$$\begin{Bmatrix} x & u & z \\ \bullet & + & y \end{Bmatrix}.$$

The theory predicts that a conditional whose utility grid fits this configuration triggers the inference that agent x will take action p, by application of the folk axiom of limited altruism.

For our current purpose the most important aspect of the theory is its utility grid tool, because the permutation of parameters within utility grids allows the generation of a vast number of potential consequential arguments, over and beyond previously identified argument forms. To illustrate, let us start from a Threat utility grid:

[2]A reviewer raised the important question of why we should need a new language (utility grids and folk axioms) for analysing utility conditionals, instead of using the tools already provided by game theory. This question arguably deserves its own paper, and I will not go any further here than saying that game theory is primarily a normative model of what rational people should do, whereas the theory of utility conditionals is primarily a descriptive model of what lay people think others like them will do.

$$\left\{\begin{matrix} h & - & s \\ s & - & h \end{matrix}\right\}.$$

This grid could for example reflect the statement:

(11) If you hurt me, I'll hurt you.

Now let us modify the value of one parameter in the grid:

$$\left\{\begin{matrix} e & - & s \\ s & - & h \end{matrix}\right\}.$$

Now the grid would, for example, reflect the statement:

(12) If she hurts me, I'll hurt you.

Most of us would probably agree to call (12) a threat, although its structure does not fit the definition of threats that we have encountered earlier in this article. It is quite possible that findings obtained with classic threats will generalise to statements such as (12), but what if we change one more parameter in the grid? For example:

$$\left\{\begin{matrix} e & - & h \\ s & - & h \end{matrix}\right\}.$$

This new grid reflects a statement such as:

(13) If she hurts you, I'll hurt you.

Now we are definitely in unknown territory. No easy name comes to mind as to what speech act (13) might be, and no immediate conclusion comes to mind for which (13) would be an argument. Note that other grids correspond to statements for which we do not have an easy name, but that we nevertheless understand as consequential arguments, for example:

$$\left\{\begin{matrix} s & + & h \\ e & - & s \end{matrix}\right\}$$

That is:

(14) If I help you, she'll hurt me.

This statement is not a promise, nor a warning, nor a slippery slope argument, nor anything similar, but it does seem to perform as a consequential argument for the conclusion that the speaker will not help the listener. As this example illustrates, we can capitalise on the utility grid representation to generate new forms of consequential arguments, and therefore attain a broader perspective on consequential argumentation. What is more, we have the possibility of testing the robustness of our models of consequential argumentation by engaging in random sampling of stimuli within the whole set of arguments defined by the combinatorial of utility grids parameters. To illustrate this novel approach, the following section reports an experiment where four utility grids were randomly selected, given minimal linguistic content to turn them into arguments, and presented to participants in order to assess the conclusions these arguments were perceived as supporting.

METHOD

Four utility grids were randomly generated using the procedure described in Figure 1. Only a few liberties were taken with respect to pure randomness: (a) the neutral agent ω was not allowed in the top-left cell, (b) y (resp., y') was forced to $\mathcal{A}$ when u (resp., u') was 0, and (c) if a grid corresponded to a

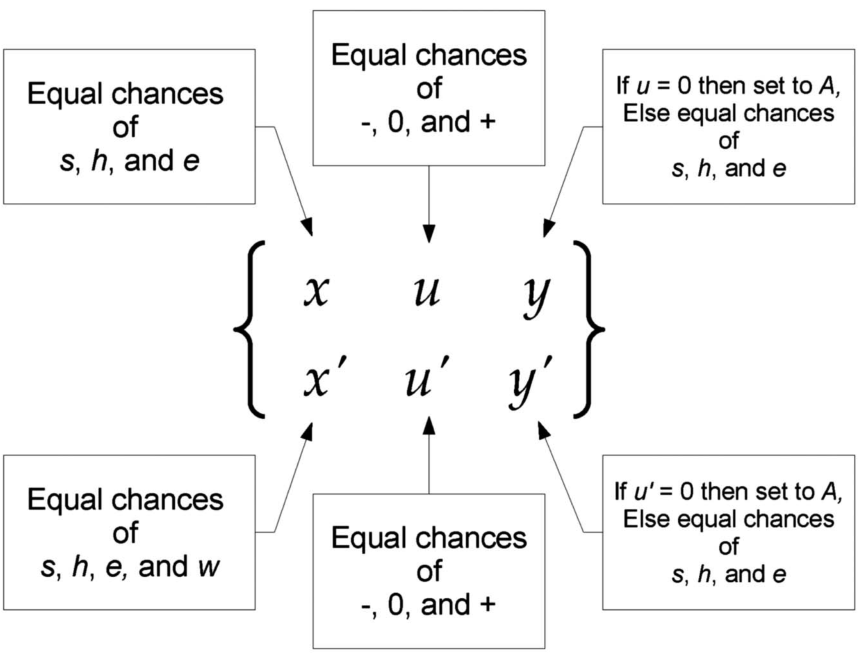

Figure 1. Procedure used to generate the random utility grids.

previously investigated argument form, it was discarded and generated again.

Rule (a) was adopted because grids where the x agent is not human are a very special case that would not be properly addressed here. Rule (b) is a convention of the utility grid representation. Rule (c) was adopted to avoid redundancy with prior results. It was applied once, when a grid turned out to correspond to a consequential conditional in the sense of Bonnefon and Hilton (2004).

The four resulting grids are shown in Table 1, together with their linguistic interpretation. In order to construct this linguistic interpretation, some minimal conventions were adopted: Third parties were given a first name, an action with a null utility was described as "doing this", an action with positive utility was described as "helping", and an action with a negative utility was describe as "hurting".

The four statements in Table 1 were presented to 39 students at the University of Toulouse (5 men, mean age = 22.6, $SD = 2.9$), in two counterbalanced orders. After each statement, participants indicated whether the x agent would take the p action, e.g., "According to you, is Sally going to hurt the speaker?", using an 11-point scale graduated from -5 (*Probably not*) to $+5$ (*Probably yes*). The full introduction to the experiment was as follows (translated from French):

> We are going to present you with four situations, which are very briefly described. In each situation a different character makes a statement, for example, "If I eat oysters, I will be very sick." We then ask you whether you expect someone to do something, for example, "Do you expect the speaker to eat oysters?" Obviously, there is no right or wrong answer, we just want to know what comes to your mind. To respond, please tick on the scale that goes from -5 to $+5$, where zero means you have no idea, neither in one direction nor the other.

RESULTS AND DISCUSSION

The four statements were analysed separately, in order to assess which statements if any led to responses that deviated significantly from the mid-point of the scale. This was the case for two statements.

TABLE 1
The four utility grids generated for the experiment, together with their minimal linguistic interpretation

$\begin{Bmatrix} e_1 & 0 & \mathcal{A} \\ e_2 & + & s \end{Bmatrix}$	If Marie does this, Alain will help me	$\begin{Bmatrix} s & 0 & \mathcal{A} \\ h & - & h \end{Bmatrix}$	If I do this, you will hurt yourself
$\begin{Bmatrix} h & - & s \\ s & + & e \end{Bmatrix}$	If you hurt me, I will help Luis	$\begin{Bmatrix} e & + & s \\ s & - & h \end{Bmatrix}$	If Sally helps me, I will hurt you

When presented with the statement "If Marie does this, Alain will help me", and asked whether Marie would do this, the participants' average response was +1.0 ($MSE = 0.3$, $d = 0.5$, 95% confidence interval: +0.4, +1.7). This was a significant deviation from zero, $t(38) = 3.4$, $p < .01$ (p-values are two-tailed). For "If I do this, you will hurt yourself", the average response was −2.7 ($MSE = 0.3$, $d = 1.4$, 95% confidence interval: −3.3, −2.1), which is a significant deviation from zero, $t(38) = -8.6$, $p < .001$.

The two other statements yielded average responses that did not significantly deviate from the mid-point of the scale. For "If you hurt me I will help Luis", the average response was −0.3 ($MSE = 0.5$, $d = 0.1$, 95% confidence interval: −1.2, +0.6), $t(38) = -0.6$, ns. For "If Sally helps me I will hurt you", the average response was −0.5 ($MSE = 0.4$, $d = 0.2$, 95% confidence interval: −1.3, +0.5), $t(38) = -1.2$, ns.

While the current experiment was not conducted as a test of the theory of utility conditionals, we can observe that the results are predicted by the theory. The folk axiom of limited altruism, applied to the utility grids of "If Marie does this, Alain will help me" and "If I do this you will hurt yourself", predicts the conclusion that Mary will do this, and that the speaker will not, respectively.

The utility grid of "If you hurt me I will help Luis" is more complex because it triggers two conflicting applications of limited altruism. Because p would both directly hurt an agent and indirectly benefit another agent, limited altruism points both to the conclusion that the listener will help and that she will not. In such a case of conflict the theory predicts that reasoners will not draw any conclusion, which corresponds here to choosing the mid-point of the scale. A similar account can be given of "If Sally helps me I will hurt you". To be precise, the theory also considered the possibility of a *primum non nocere* principle (i.e., "above all, do no harm"), to the effect that such conflicts might lead to the conclusion that the action would not be taken. The fact that the current data are skewed towards the negative end of the scale suggests that this possibility must not be discarded yet.

The findings illustrate the heuristic value of the utility grid representation for generating new forms of consequential arguments. A statement such as "If I do this, you will hurt yourself" is clearly perceived as a consequential argument, in spite of not being easily sortable in any of the categories that formed the bulk of previous research. The random sampling of stimuli has an additional interest, which is to generate examples whose interpretation is less straightforward than standard promises or threats. For example, participants seemed to have trouble interpreting the statement "If you hurt me I will help Luis", for which the distribution of responses was essentially flat. It could be that some conditionals, such as this one, have a Necker-like quality in terms of the utility grid that they reflect, or the speech act that they perform. It is quite possible that some participants reinterpreted the statement as a threat (i.e., the listener does not want Luis to be helped), whereas some other participants reinterpreted the statement as a promise (i.e., the speaker enjoys being hurt). This in turn leaves open the possibility that reasoners rearrange many non-standard utility grids into the classic grids explored by previous research. If this were to turn out to be true, the prospect of reaching a complete understanding of consequential arguments would be far more plausible than expected.

CONCLUSION

Most of the empirical research on consequential arguments has focused on slippery slopes and inducive speech acts. While these subsets have intrinsic interest, and have deserved specific investigations, the scope of research on consequential arguments has been limited by the lack of a systematic method to address the structural diversity of these arguments. I have suggested using the theory of utility conditionals and its utility grid representation as a way to generate potential consequential arguments.

It is important to note that, just as the theory of utility conditionals is not in competition with the approaches discussed in the introduction to this article, so the random sampling approach advocated in this article is not in competition with the focused investigations conducted in other research programmes. The theory of utility conditionals focuses on the generation of default inferences, where other approaches to argumentation focus on the provision and evaluation of reasons for a claim. As a consequence, the theory of utility conditional has little to say yet about the graded strength or persuasiveness of an argument, in contrast to other approaches such as the Bayesian model of consequentialist arguments. Focused investigations of argument strength are just as essential to the study of consequentialist arguments as the random sampling approach I have illustrated in this article. Although this illustrative experiment was limited to four novel arguments, it is expected that further use of this method will bring more diversity to experimental research on consequential arguments, and more robustness to models of consequential argumentation.

REFERENCES

Amgoud, L., Bonnefon, J. F., & Prade, H. (2005). An argumentation-based approach for multiple criteria decision. *Lecture Notes in Computer Science*, *3571*, 269–280.

Amgoud, L., Bonnefon, J. F., & Prade, H. (2007). The logical handling of threats, rewards, tips, and warnings. *Lecture Notes in Artificial Intelligence*, *4724*, 235–246.

Beller, S., Bender, A., & Kuhnmünch, G. (2005). Understanding conditional promises and threats. *Thinking & Reasoning*, *11*, 209–238.

Bonnefon, J. F. (2004). Reinstatement, floating conclusions, and the credulity of mental model reasoning. *Cognitive Science*, *28*, 621–631.
Bonnefon, J. F. (2009). A theory of utility conditionals: Paralogical reasoning from decision-theoretic leakage. *Psychological Review*, *116*, 888–907.
Bonnefon, J. F., Dubois, D., Fargier, H., & Leblois, S. (2008). Qualitative heuristics for balancing the pros and cons. *Theory and Decision*, *65*, 71–95.
Bonnefon, J. F., Girotto, V., & Legrenzi, P. (2012). The psychology of reasoning about preferences and unconsequential decisions. *Synthese*, *187*, 27–41.
Bonnefon, J. F., & Hilton, D. J. (2004). Consequential conditionals: Invited and suppressed inferences from valued outcomes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 28–37.
Corner, A., & Hahn, U. (2007). Evaluating the meta-slope: Is there a slippery slope argument against slippery slope arguments? *Argumentation*, *21*, 349–359.
Corner, A., & Hahn, U. (2009). Evaluating science arguments: Evidence, uncertainty and argument strength. *Journal of Experimental Psychology: Applied*, *15*, 199–212.
Corner, A., Hahn, U., & Oaksford, M. (2011). The psychological mechanism of the slippery slope argument. *Journal of Memory and Language*, *64*, 153–170.
Dubois, D., Fargier, H., & Bonnefon, J. F. (2008). On the qualitative comparison of decisions having positive and negative features. *Journal of Artificial Intelligence Research*, *32*, 385–417.
Evans, J. St. B. T. (2005). The social and communicative function of conditional statements. *Mind & Society*, *4*, 97–113.
Evans, J. St. B. T., Neilens, H., Handley, S. J., & Over, D. E. (2008). When can we say 'if'? *Cognition*, *108*, 100–116.
Evans, J. St. B. T., & Twyman-Musgrove, J. (1998). Conditional reasoning with inducements and advice. *Cognition*, *69*, B11–B16.
Hahn, U., & Oaksford, M. (2006). A Bayesian approach to informal argument fallacies. *Synthese*, *152*, 207–236.
Hahn, U., & Oaksford, M. (2007). The rationality of informal argumentation: A Bayesian approach to argument fallacies. *Psychological Review*, *114*, 704–732.
Haigh, M., Stewart, A. J., Wood, J. S., & Connell, L. (2011). Conditional advice and inducements: Are readers sensitive to implicit speech acts during comprehension? *Acta Psychologica*, *136*, 419–424.
Hoeken, H., Timmers, R., & Schellens, P. J. (2012 this issue). Arguing about desirable consequences: What constitutes a convincing argument? *Thinking & Reasoning*, *18,* 394–416.
Legrenzi, P., Politzer, G., & Girotto, V. (1996). Contract proposals: A sketch of a grammar. *Theory & Psychology*, *6*, 247–265.
López-Rousseau, A., & Ketelaar, T. (2006). Juliet: If they do see thee, they will murder thee: A satisficing algorithm for pragmatic conditionals. *Mind & Society*, *5*, 71–77.
Miller, D. T. (1999). The norm of self-interest. *American Psychologist*, *54*, 1053–1060.
Ohm, E., & Thompson, V. (2004). Everyday reasoning with inducements and advice. *Thinking & Reasoning*, *10*, 241–272.
Ohm, E., & Thompson, V. (2006). Conditional probability and pragmatic conditionals: Dissociating truth and effectiveness. *Thinking & Reasoning*, *12*, 257–280.
Rahwan, I., Madakkatel, M. I., Bonnefon, J. F., Awan, R. N., & Abdallah, S. (2010). Behavioral experiments for assessing the abstract argumentation semantics of reinstatement. *Cognitive Science*, *34*, 1483–1502.
Smedslund, J. (1997). *The structure of psychological common sense*. Mahwah, NJ: Lawrence Erlbaum Associates Inc.
Thompson, V. A., & Evans, J. St. B. T. (2012 this issue). Belief bias in informal reasoning. *Thinking & Reasoning*, *18,* 278–310.

Thompson, V. A., Evans, J. St. B. T., & Handley, S. J. (2005). Persuading and dissuading by conditional argument. *Journal of Memory and Language*, *53*, 238–257.

Verbrugge, S., Dieussaert, K., Schaeken, W., & Van Belle, W. (2004). Promise is debt, threat another matter: The effect of credibility on the interpretation of conditional promises and threats. *Canadian Journal of Experimental Psychology*, *58*, 106–112.

Walton, D. (1999). Historical origins of argumentum ad consequentiam. *Argumentation*, *13*, 251–264.