

# Northumbria Research Link

Citation: Bonnefon, Jean-François, Haigh, Matthew and Stewart, Andrew (2013) Utility templates for the interpretation of conditional statements. *Journal of Memory and Language*, 68 (4). pp. 350-361. ISSN 0749-596X

Published by: Elsevier

URL: <http://dx.doi.org/10.1016/j.jml.2013.01.002>  
<<http://dx.doi.org/10.1016/j.jml.2013.01.002>>

This version was downloaded from Northumbria Research Link:  
<https://nrl.northumbria.ac.uk/id/eprint/13554/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)

Running head: UTILITY TEMPLATES

Utility Templates for the Interpretation of Conditional Statements

Jean-François Bonnefon

Centre National de la Recherche Scientifique

Matthew Haigh and Andrew J. Stewart

University of Manchester

Address for correspondence

Jean-François Bonnefon, CLLE

Maison de la recherche, 5 allées A. Machado, 31058 Toulouse Cedex 9, France

Telephone: 33 5 61 50 36 01

E-mail: [bonnefon@univ-tlse2.fr](mailto:bonnefon@univ-tlse2.fr)

### **Abstract**

People often use conditional statements to describe configurations of agents, actions and valued consequences. The article postulates the existence of utility templates, a special subset of these configurations that exert strong constraints on interpretation. A completion survey identified four potential templates, and four experiments documented their characteristic effects: When a situation is close enough to a template, people interpret ambiguous information or reinterpret current information in such a way that their understanding of the situation fits the template. A process explanation of these effects is considered, which allows for the principled generation of other templates, and offers a possible reformulation of the findings within the framework of relevance theory.

### Utility Templates for the Interpretation of Conditional Statements

People use conditional sentences when describing actions, their preconditions, and their consequences. Oftentimes, these actions and consequences matter, in the sense that they have value or utility to various agents:

- (1)
  - a. If she praises him, he will support her.
  - b. If I hire you, I do you a favor.
  - c. If she harasses him, she will be fired.
  - d. If you leave me, I will be crushed.

The conditionals above feature various arrangements of agents, actions, and values of these actions. For example, Conditional (1-a) says that an agent  $x$  will do something positive for another agent  $y$ , provided that  $y$  does something positive for  $x$ . Many other arrangements of agents, actions and values are possible, of which Conditionals (1-a-d) offer but a small sample. Our contention in this article, though, is that some of these arrangements have the special status of *utility templates* that guide and constrain interpretation. Utility templates are scripts, schemas that help people make sense of situations where various agents perform valued actions. More precisely, we argue that whenever it is possible, people will interpret or re-interpret a conditional sentence in order to make it coincide with one of their utility templates. Consider for example Conditional (2), in which the then-clause features a non-verb whose meaning has to be inferred:

- (2) If Alice supports Bob, Bob will yorb her.

We will argue that such a conditional prompts people to activate a Social Contract utility template, from which they infer that Alice likes to be yorbed. If, on the contrary, they are

informed that Alice dislikes being yorbed, the same utility template should lead them to infer that Bob dislikes being supported by Alice.

In order to flesh out our proposal, we will take advantage of the theory of utility conditionals (Bonnefon, 2009), which offers a convenient notation to represent the arrangement of agents, actions and values packed in a conditional sentence. After we have introduced this notation, we consider the most natural candidate as a utility template, the Social Contract illustrated in (1-a). We then report data from a sentence completion survey that helped us identify three other candidates, the Tautology, Justice, and Unpacking templates, illustrated in (1-bcd). Finally, we report the findings of four experiments investigating whether and to which extent these templates constrain the interpretation of conditional sentences featuring non-verbs in their if- or then-clause.

The theory of utility conditionals expanded prior content-specific approaches to conditionals by providing a general classification schemes for all utility conditionals. Previous approaches were usually limited to specific types of utility conditionals, such as promises, permissions, or precaution rules. The strength of the theory of utility conditionals was to show that all these specific conditional patterns were spots within a greater territory, and to offer a systematic map of this territory. This systematic effort pointed to a daunting number of utility conditionals: Hundreds, possibly thousands of utility conditionals could now be defined and studied. Here, we consider the possibility that some of these utility conditionals have the special status of utility templates. Importantly, this means that entire parts of the map (and the dozens of utility grids that they encompass) can be subsumed under a single template. In sum, after the expansion phase that corresponded to the theory of utility conditionals, this article offers a contraction phase. While it could seem intractable to explore the hundreds of patterns identified by the theory of utility conditionals, the task will be much easier if large subsets of patterns can be subsumed under a tractable number of templates.

### Utility Conditionals

The theory of utility conditionals (Bonnefon, 2009) was developed to predict the inferences that people draw from conditional sentences featuring valued actions and consequences (Bonnefon, Girotto, & Legrenzi, 2012; Bonnefon & Hilton, 2004; Bonnefon & Sloman, in press; Corner, Hahn, & Oaksford, 2011; Evans, Neilens, Handley, & Over, 2008; Haigh, Stewart, Wood, & Connell, 2011; Ohm & Thompson, 2004; Thompson, Evans, & Handley, 2005). For our current purpose, we are mostly interested in the systematic notation scheme that the theory affords. The theory assumes that “if  $p$ , then  $q$ ” sentences can be unpacked as “if agent  $x$  does  $p$  with utility  $u$  to agent  $y$ , then agent  $x'$  will do  $q$  with utility  $u'$  to agent  $y'$ ”. This information is represented in compact form in the *utility grid* of the conditional:

$$\left\{ \begin{array}{ccc} x & u & y \\ x' & u' & y' \end{array} \right\}.$$

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.

For convenience, the agent who states the conditional is noted  $s$  (for ‘speaker’), the agent at whom the conditional is directed is noted  $h$  (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  $\omega$ . The agent  $\omega$  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 the whole set of agents  $\mathcal{A}$  as a target.<sup>1</sup>

The critical advantage of this notation is that it can characterize familiar speech acts (e.g., promises, warnings) as well as statements for which we do not have a convenient label, but whose utility grid is nevertheless unique (Bonneton, 2012). Conditional warnings provide an example of a familiar speech act. They are typically defined as if-then statements whose if-clause features an action of the hearer and whose then-clause features a state of the world that is undesirable to the hearer (Amgoud, Bonneton, & Prade, 2007; Evans, 2005; López-Rousseau, Diesendruck, & Benozio, 2011; López-Rousseau & Ketelaar, 2006), for example:

- (3)
- a. If you click here, your computer will be infected by a virus.
  - b. If you wait any longer, you will be late for your plane.
  - c. If you go partying, you will fail your test tomorrow.
  - d. If you denounce me, you will have an accident.

All these warnings have a utility grid of the form:

$$\left\{ \begin{array}{ccc} h & \bullet & \bullet \\ \omega & - & h \end{array} \right\},$$

where the black dot can take various values reflecting different subclasses of warnings, some of which are illustrated in (3).

Beyond characterizing familiar speech acts such as warnings, the permutation of parameters in the utility grid allows one to identify brand new subclasses of conditionals.

Consider for example the grid:

$$\begin{pmatrix} s & + & h \\ e & - & s \end{pmatrix},$$

which would reflect a statement such as:

- (4) If I help you, she will hurt me.

This statement is not a promise, nor a warning, nor anything similar. The appeal of utility grids, and of the theory of utility conditionals more generally, is that they allow one to predict the inferences prompted by Conditional (4) (e.g., the listener believes that the speaker should help, but the speaker will not help), without the need to precisely define what speech act it performs. The grid notation thus allows research to move beyond the limited subset of lexicalized conditional speech acts, and provides a systematic map of the vast unknown territory of utility conditionals.

In the rest of this article, we will take advantage of the notational power of utility grids in order to identify what we call *utility templates*, that is, special grid configurations that guide and constrain the interpretation of conditional statements. We begin with considering the most obvious candidate as a utility template (the Social Contract), before we move on to other potential templates.

### The Social Contract Template

Most generally, a social contract describes a situation where an agent  $x$  is entitled to a benefit  $q$  provided by an agent  $y$ , on the condition that agent  $x$  satisfies a requirement  $p$  set by agent  $y$ . Conditional sentences are especially apt to express social contracts:

- (5) a. If you wash my car, I'll let you borrow it tonight.  
 b. If Alice supports Bob, Bob will reward her.  
 c. If they apologize to me, I'll drop the charges against them.

Conditionals (5-a-c) all have a utility grid of the form:

$$\begin{Bmatrix} x & + & y \\ y & + & x \end{Bmatrix},$$

where  $x$  and  $y$  take appropriate values, for example “Alice” and “Bob”, respectively, in Conditional (5-b). Note though that some social contracts can stipulate the retaliatory measures that  $y$  may take when  $x$  fails to meet some requirement deemed desirable by  $y$ :

- (6)
- a. If you cheat on me, I will leave you.
  - b. If Alice demotes Bob, Bob will sue her.
  - c. If they show up late, I'll deny them entrance.

All these conditionals have the same grid as displayed above, with minus signs instead of plus signs. The general form of a social contract can thus be expressed with the following grid, where  $u \neq 0$ :

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

A considerable amount of scholarship suggests that the situations captured by this grid (and the sentences that describe them) have particular cognitive salience. Legrenzi, G., and Girotto (1996) called it the canonical form of contract proposals, about which, they remarked, people could reason with remarkable accuracy (Politzer & Nguyen-Xuan, 1992). Various content-dependent approaches to conditional inferences offered explanations for this accuracy, in terms of repeated exposure leading to the acquisition of a mental schema (Cheng & Holyoak, 1985), or as the result of an adaptation to the evolutionary problem of cheater detection (Cosmides, Barrett, & Tooby, 2010), which endowed us with an innate Darwinian algorithm for processing social contracts. What is clear is that people have an exquisite

sensitivity to the utility structure underlying social contracts, which manifests itself both in reading (Haigh et al., 2011) and reasoning experiments (e.g. Hilton, Kimmelman, & Bonnefon, 2005; Perham & Oaksford, 2005).

In line with all this evidence, we suggest that the grid above serves as a template for interpreting conditional sentences which feature actions and consequences that have utility for various agents. We expect that this template will take precedence when a sentence is ambiguous, in the sense that it could reflect several grids, one of which is the Social Contract template. Consider for example the following sentence, which features an unknown verb in the then-clause of the conditional:

(7) If Alice supports Bob, then Bob will yorb her.

This sentence is captured by the following grid, where the question mark signals that being yorbed has unknown utility to Alice:

$$\begin{pmatrix} a & + & b \\ b & ? & a \end{pmatrix}.$$

Being yorbed could be good or bad to Alice, or she could be indifferent to being yorbed. Only the first of these interpretations, though, would turn the grid into a Social Contract template. Our prediction is that people will apply the Social Contract template to the sentence, and consider that Alice likes being yorbed.

What now if people are told for a fact that Alice dislikes being yorbed? This piece of information would result in a grid that would violate the Social Contract template:

$$\begin{pmatrix} a & + & b \\ b & - & a \end{pmatrix}.$$

If we are correct that the Social Contract template exerts constraints on interpretation, then people should be led to revise their interpretation in a way that both fits the template and the piece of information they received. The only route thereto is to turn the plus sign in the first row into the minus sign; that is, to consider that Bob, for whatever reason, dislikes being supported by Alice. This is, we predict, the route that people will take.

The experiments we report in this article test these two basic predictions about utility templates: that they guide the interpretation of unknown verbs, and that they lead to revise the interpretation of known verbs whose typical valence conflicts with a template. Our predictions go beyond previous research along two dimensions. First, we make predictions about how individuals *assign* utility to the clauses of a conditional, rather than how individuals infer conclusions *once* they have identified these utilities. Second, our experiments are not limited to the Social Contract template (or other lexicalized speech acts such as permissions or precaution rules): They also address three other templates, that we tentatively identified through a sentence completion survey described in the next section. Indeed, the fact that the Social Contract featured prominently in previous, content-specific approaches to conditionals suggests that it could be a template – but not that it is the only one, and not even that it is the strongest or most frequent one. Many scholars observed that individuals could quickly detect Social Contracts, and accurately reason from them, and we accordingly expect Social Contracts to display the characteristics we expect from utility templates. But we also expect that other conditionals, for which we might not have a convenient lexical label, will also display these characteristics. Some of these conditionals might have escaped the attention of previous research only because they were not lexicalized speech acts. The theory of utility conditionals, with its ability to represent all utility patterns, will allow us to identify other template candidates.

### Other Templates

In order to identify potential templates, we conducted a survey in which 30 participants (students at the University of Manchester, 18 women, mean age 27.5 years, SD 8.6 years) were asked to complete 30 conditional fragments such as:

(8) If I give you a job then I . . .

These 30 fragments reflected all the unique utility grid permutations with two blank cells (bottom-center and bottom-right), limited to the interactions between agents  $s$  and  $h$ , or between agents  $e_1$  and  $e_2$ . The 30 fragments are listed in Appendix A.

Participants were asked to complete the 30 fragments so that they formed grammatically correct English phrases. This provided a total of 900 complete conditionals. The items were presented visually, in a different random order for each participant. All participants were native English speakers and completed the task in individual testing booths following an unrelated experiment. They were each compensated £5.

The first and second authors jointly coded participants' responses in order to assign a utility grid to each of the 900 completed conditionals, using the coding scheme outlined by Bonnefon (2009). A substantial number of fragments (35%) were completed with consequents that had no clear utility to any agent, but the remaining fragments did point to a small number of potential utility templates. These potential templates are displayed in Table 1.

As could be expected, the Social Contract grid emerged from the survey, accounting for 11% of completions. Here are some typical examples of completions fitting the Social Contract grid:

- (9)
- a. If Colin helps Laura, then [she must help him too].
  - b. If I give you a job, then [you must work hard for me].
  - c. If Robert insults Joanne, then [she will probably slap him].

- d. If you hurt me, then [I will hurt you back].

Interestingly though, another grid accounted for a greater 13% of completions. We called it the Justice grid, as it seemed to encapsulate the idea that good things happen to people who do good deeds, whereas bad things happen to people who do bad deeds. Here are some typical examples of completions fitting the Justice grid:

- (10) a. If John helps Sarah, then [he will be rewarded].
- b. If I give you a job, then [I will feel good about myself].
- c. If Brian insults Mandy, then [he will get told off].
- d. If you hurt me, then [you will feel guilty].

A third grid accounted for an additional 10% of completions, which we called the Unpacking grid. Consider first some typical examples fitting that grid:

- (11) a. If Colin helps Laura, then [she will get her problems sorted out].
- b. If I give you a job, then [you will have money].
- c. If Robert insults Joanne, then [she will be upset].
- d. If you hurt me, then [I will not be happy].

What people seem to do in these completions is to unpack the utility that was already implicit in the antecedent. For example, having one's problems sorted out is the very reason why most people find positive utility in being helped; and getting to earn money is the very reason why people rejoice about finding a job.

No other grid seemed to account for enough completions to be considered a potential template. One grid, though, attracted our attention despite of its accounting for only 2% of completions. We call it the Tautology grid because its bottom line is merely a repetition of its top line. For example:

- (12)
- a. If John helps Sarah, then [he is doing her a favor].
  - b. If I give you a job then [I am helping you].
  - c. If Brian insults Mandy, then [he will upset her].
  - d. If you hurt me, then [you will really upset me].

This grid was fairly rare in the completion survey, but it still struck us as a potential utility template, because of its structural simplicity and its closeness to the Unpacking grid. We thus decided to tentatively add it to the group of potential templates tested in the studies that we report in this article.

### Study 1

This first study was designed to investigate whether the utility templates identified in the completion survey would guide the interpretation of nonverbs whose valence was unknown to participants.

#### *Method*

Sixty participants from the University of Manchester (54 women) completed the task for course credit. None of them had taken part in the completion survey. The task was administered online using survey software. Participants were presented with 16 conditional sentences that contained nonverbs in both their clauses, for example:

- (13)
- a. If Graham yorbs Sarah, then Sarah will naft Graham;
  - b. If Adam kawps Sue, then Adam will sork Sue;
  - c. If Lisa murbs Ian, then she will be tymped;
  - d. If Peter peens Claire, then she will be dested.

Each non-verb was four characters long and did not resemble any real word that could feasibly be used in its place. The nonverbs were selected from the ARC nonword database (Rastle,

Harrington, & Coltheart, 2002). Among the 16 conditionals, four had the structure of a Social Contract template as in (13-a), four had the structure of an Unpacking template as in (13-b), four had the structure of a Justice template as in (13-c), and four had the structure of Tautology template as in (13-d). To clarify, what we mean for a conditional to ‘have the structure of’ a Template is that the four non-utility slots in the grid of the conditional are consistent with the Template. Our test then consists in providing information about the utility of one nonverb, and measuring whether participants assign to the other nonverb the utility predicted by the Template (that is, a utility of the same sign).

Thus, participants received information about the utility of one nonverb (henceforth, the companion nonverb), and had to rate the utility of the other nonverb (henceforth, the target nonverb). The purported utility of the companion nonverb could be positive or negative, leading to a  $4 \times 2$  design, manipulating utility template and valence of companion nonverb, with two statements in each condition in order to counterbalance the position of the companion nonverb (in the if-clause or in the then-clause). A typical item read:

If Lisa murders Ian, then she will be typed. Lisa dislikes to be typed. How much would Ian like to be murdered, on other occasions than this one?

The question always specified “on other occasions than this one” to capture the intrinsic utility of the target nonverb, rather than its second-order utility. For example, in the sentence above, Lisa dislikes to be typed, so Ian could see second-order negative utility in events that would lead, in this specific occasion, to the typing of Lisa. We are not interested in this second-order utility, but only in participants’ perception of how much Ian likes to be murdered in general. Participants answered on a 7-point scale anchored at  $-3$  (*not at all*) and  $+3$  (*very much*), where zero indicated indifference. We predicted that the utility assigned to the target nonverb would be of the same sign as that of its companion nonverb.

## Results

Table 2 displays the utility ratings of the target nonverbs, as a function of the valence of the companion nonverbs, and the type of utility template. A preliminary analysis showed that the position of the target nonverb had no detectable effect – this factor is not represented in Table 2. Quite evidently, and in line with our expectations, participants assigned positive utility to target nonverbs when their companion nonverb were of purportedly positive valence; and they assigned negative utility to target nonverbs when their companion nonverb were of purportedly negative valence.

We conducted a  $2 \times 4$  repeated measures analysis of variance on the utility ratings, with the valence of the companion nonverb and the utility template as repeated factors. This analysis detected a main effect of the companion nonverb valence,  $F(1, 59) = 230.8$ ,  $p < .001$ ,  $\eta_p^2 = .80$ , and an interaction between valence and utility template,  $F(3, 57) = 4.0$ ,  $p = .012$ ,  $\eta_p^2 = .17$ . The main effect of utility template was non-significant ( $F < 1$ ).

To determine the variables driving this interaction we collapsed the data over valence by calculating the mean of the absolute target nonverb ratings for each of the four templates. Pairwise comparisons revealed that absolute ratings for Unpacking templates were greater than for Tautology templates,  $t(59) = 2.027$ ,  $p = .047$ , and Justice templates  $t(59) = 2.997$ ,  $p = .004$ . Ratings for Social Contract templates were also greater than for Justice templates  $t(59) = 2.223$ ,  $p = .03$ . All other pairwise comparisons were non significant ( $p > .05$ ).

Although the effect of the companion nonverb was huge for all templates, it was larger for the Unpacking and Contract templates than for the Justice template. It might be that the Unpacking and Contract templates impose stronger constraints on interpretation, but we will postpone this discussion until after we have reported our series of experiments, as these other experiments will help us to assess the robustness of this result. We now report Study 2, in which we tested our second main prediction: Participants revise the valence they attribute to a known verb, when it is paired with a nonverb of conflicting valence within a utility template.

## Study 2

### *Method*

Fifty six participants from the University of Manchester population (49 women) completed the task for course credit. The task was administered online, using survey software. None of these participants took part in the completion survey or in Study 1.

Participants were presented with 16 conditional sentences in a  $4 \times 2 \times 2$  design (Template by Valence by Congruency). There were four sentences for each template that contained one nonverb and one real verb, which could be either “help” (positive utility) or ‘hurt’ (negative utility). In the control condition, participants simply rated the perceived utility of the verb, for example:

If Alice hurts Bob, then Alice is zimmered. How much does Bob like to be hurt?

Participants responded on a 7-point scale anchored at  $-3$  (*not at all*) and  $+3$  (*very much*), where zero indicated indifference. In the conflict condition, participants were provided with an additional piece of information about the utility of the nonverb. This utility was in conflict with what could be expected from the template. For example:

If Alice hurts Bob, then Alice is zimmered. Alice likes to be zimmered. How much does Bob like to be hurt?

We predicted that participants would revise their interpretation in order to reconcile their utility template and the piece of information they received. That is, we predicted that they would change the utility they assigned to the real verb, assigning negative utility to being helped and positive utility to being hurt.

### *Results*

Figure 1 displays the utility assigned to being helped or being hurt, for each utility template, in the control and conflict conditions. The visual inspection of Figure 1 suggests that

our predictions were correct: In the conflict condition, participants assigned negative utility to being helped (whereas they did not in the control condition), and they assigned positive utility to being hurt (whereas they did not in the control condition). This is confirmed by a  $4 \times 2 \times 2$  ANOVA which revealed a significant three way interaction between utility template, valence and congruency,  $F(3, 53) = 13.2, p < .001, \eta_p^2 = .43$ .

To examine the factors driving the three way interaction we performed separate  $4 \times 2$  (Template by Congruency) ANOVAs for each level of Valence (i.e., Help and Hurt). When the verb was positive (i.e., help) the ANOVA revealed main effects of utility template  $F(1, 53) = 3.3, p = .03, \eta_p^2 = .16$ , and congruency  $F(1, 55) = 71.7, p < .001, \eta_p^2 = .57$ ; as well as a significant interaction between these two variables  $F(3, 53) = 8.3, p < .001, \eta_p^2 = .32$ . Pairwise comparisons showed that the congruency effect was obtained for all templates: Unpacking,  $t(55) = 8.61, p < .001$ ; Contract,  $t(55) = 3.71, p < .001$ ; Tautology,  $t(55) = 5.52, p < .001$ ; and Justice,  $t(55) = 2.44, p < .02$ .

When the verb was negative (i.e., hurt) the ANOVA also revealed main effects of the utility template,  $F(3, 53) = 14.2, p < .001, \eta_p^2 = .45$ , and congruency,  $F(1, 55) = 56.0, p < .001, \eta_p^2 = .51$ , as well as a significant interaction between these two variables,  $F(3, 53) = 7.8, p < .001, \eta_p^2 = .31$ . Pairwise comparisons showed that the congruency effect was not obtained for the Justice template,  $t(55) = 1.34, p = .19$ , but was obtained for the other three templates: Unpacking,  $t(55) = 5.86, p < .001$ ; Contract,  $t(55) = 6.66, p < .001$ ; Tautology,  $t(55) = 4.21, p < .001$ .

As in Study 1, the predicted effect was observed for all templates (albeit partially for the Justice template). The effect was especially strong for the Unpacking template, as it was already the case in Study 1. However, unlike what happened in Study 1, the effect was not stronger for the Contract Template than it was for the Tautology template. Once more, we will postpone our discussion of these results until after we have reported our other studies.

Studies 3 and 4 address two concerns about the data we have reported so far. First, the

results observed in the control condition of Study 2 are odd. As seen in Figure 1, participants tended to rate 'help' and 'hurt' as having neutral utility in the control condition. Quite likely, this behavior reflected the demand characteristics of the experiment. Participants could quickly figure out the design of our experiment, and they might have considered that they could not express an informed opinion for items in which no information was offered about the utility of the nonverb. Study 3 will clarify experimental demands by making it clear that participants should rely on their default interpretation of the verbs in the control condition.

Second, all the templates we have identified in the completion survey are such that the utility of the two clauses are the same. Thus, an alternate explanation for our results might be that participants understood their task as simply matching the utility of the two verbs in the sentence. In order to rule out his explanation, Study 4 will present participants with sentences featuring other connectives than the conditional, namely, conjunctions and disjunctions. If we are able to show that participants do not match the utility of the two verbs with these connectives, we will be in a position to rule out the matching explanation of the results we observed with conditional sentences.

### **Study 3**

#### *Methods*

Fifty five participants (51 women) from the University of Manchester population completed the task for course credit. The task was administered online, using survey software. None of these participants took part in the completion survey or in the previous studies. Participants provided 16 valence ratings in a  $4 \times 2 \times 2$  design (Template by Valence by Congruency). There were four sentences for each template that contained one nonverb and one real verb, which could be either 'help' (positive utility) or 'hurt' (negative utility). Participants first completed a control condition in which they simply rated the perceived utility of the verb. Responses were recorded on a 7-point scale anchored at  $-3$  (*not at all*) and  $+3$

(*very much*), where zero indicated indifference. They were then provided with an additional piece of information about the utility of the nonverb and asked to provide a new rating. The utility featured in this additional information was incongruent with what could be expected from the template. For example:

If Alice hurts Bob, then he will be zimmered. How much does Bob like to be hurt?

Now assume that Bob likes to be zimmered. How much does Bob now like to be hurt?

We predicted that participants would assign positive (resp., negative) valence to ‘help’ (resp., ‘hurt’) in the control condition, but that they would revise this interpretation in the conflict condition, in order to reconcile their utility template and the piece of information they received.

### *Results*

Figure 2 displays the utility assigned to being helped or being hurt, for each utility template, in the control and conflict conditions. The visual inspection of Figure 2 suggests that our predictions were correct: In the control condition, participants generally assigned positive utility to being helped, and negative utility to being hurt – but they revised their ratings in the conflict condition.

A repeated measures ANOVA revealed a significant three way interaction between utility template, valence and congruency,  $F(3, 52) = 20.04, p < .001, \eta_p^2 = .54$ . To examine the factors driving the three way interaction we performed separate 4 ANOVAs for ‘Help’ and ‘Hurt’. When the verb was positive (help) the ANOVA revealed a huge main effect of conflict,  $F(1, 54) = 87.06, p < .001, \eta_p^2 = .62$ , and an interaction between template and conflict,  $F(3, 52) = 9.62, p < .001, \eta_p^2 = .36$ , but no main effect of template,  $F(3, 52) = 2.54, p > .05, \eta_p^2 = .13$ . Pairwise comparisons showed that the conflict effect was obtained for Unpacking,  $t(54) = 8.67, p < .001$ ; for Tautology,  $t(54) = 9.39, p < .001$ ; and Justice,  $t(54) = 4.52, p < .001$ . The effect was only marginal for Contract templates,  $t(54) = 1.80, p = .077$ .

When the verb was negative (hurt) the ANOVA revealed a huge main effect of conflict,  $F(1, 54) = 173.52, p < .001, \eta_p^2 = .76$ , a main effect of template,  $F(3, 52) = 8.57, p < .001, \eta_p^2 = .33$ , and a significant interaction between these two variables,  $F(3, 52) = 16.69, p < .001, \eta_p^2 = .49$ . Pairwise comparisons showed that the congruency effect was obtained for all templates, Unpacking,  $t(54) = 11.12, p < .001$ ; Tautology,  $t(54) = 15.26, p < .001$ ; Justice,  $t(54) = 3.02, p < .01$ ; and Contract,  $t(54) = 6.02, p < .001$ .

In sum, Study 3 replicated the main results of Study 2 while eliciting appropriate ratings of the verbs in the control condition, assuaging a concern with the results of Study 2. Study 3 also offered an opportunity to investigate the valence by template interaction observed in the previous studies. While the Unpacking template seems to consistently elicit a stronger effect, the results for the other templates are less clear. In particular, the effect for the Contract template was weaker than that obtained in the previous studies, and the effect for the Tautology template was stronger than that obtained in the previous studies. These data make it necessary to be cautious when interpreting the valence by template interaction: only the Unpacking template seems to consistently deliver stronger results than the other templates. We will get back to this issue in the General Discussion section.

#### Study 4

Study 4 is intended to rule out a possible explanation for the effects observed in Studies 1–3. It could be that participants saw their task as simply matching the utilities of the two verbs in each sentence. To rule out this account, we conducted another study where the conditional connective was replaced with other connectives (conjunctions and disjunctions). If participants were merely matching the utilities of the two verbs, we should observe this effect with conjunctive and disjunctive statements. In contrast, we predict that the effects we observed were specific to conditional statements, and should not be observed with other connectives.

### *Methods*

Forty nine participants from the University of Manchester population (41 women) completed the task for course credit. The task was administered online, using survey software. None of these participants took part in the preceding studies.

Study 4 was almost similar to Study 2, except that it used conjunctions and disjunctions instead of conditionals. Participants were presented with conjunctive and disjunctive statements that contained one nonverb and one real verb, which could be either ‘praise’ (positive utility) or ‘hurt’ (negative utility). In the control condition participants simply rated the perceived utility of the verb. In the conflict condition participants were presented with information about the valence of the nonverb. The full design produced eight items. As an illustration, we reproduce here an item using a conjunction in the conflict condition, using the negative verb ‘hurt’:

Vicky weffed Roger and she hurt Clare. Roger likes to be weffed. How much does Clare likes to be hurt?

We predicted that participants would be influenced only by the intrinsic valence of the verb, and not by the valence assigned to the nonverb in the sentence. In other words, we predicted that participants would not match the utility of the verb to that of the nonverb.

### *Results*

A repeated measures ANOVA delivered the expected pattern of results. The analysis revealed a main effect of verb valence  $F(1, 48) = 58.4, p < .001, \eta_p^2 = .55$ , but no other main effect or interaction (all  $F_s < 1$ ). As seen in Table 3, participants appear to rely exclusively on the intrinsic valence of the verbs, and did not attempt to match it with the valence of the nonverb. Consequently, we can be confident that the results we observed in Studies 1–3 were not due to some simple heuristic of matching the utilities of the two verbs in each sentence.

## General Discussion

Conditional sentences can describe patterns of actions and their utility to various agents. There is a vast number of such patterns, some that correspond to known speech acts, as the threat in (14-a), and many that are not as easily labeled, as in (14-b):

- (14) a. If you insult me again, I'll hit you;  
b. If I help you, I won't finish my own work.

Some patterns, which we called *utility templates*, seem to describe scripted configurations of actions – schemas which can guide and constrain interpretation. Using the formal representation of utility grids (Bonnefon, 2009) and a sentence completion survey, we tentatively identified four such templates: Social Contract, Unpacking, Justice, and Tautology.

We then tested what we deemed to be the two characteristic effects of utility templates: (a) When a sentence conforms to a template but for one missing piece of information, an assumption is made that the missing piece of information is in line with the template; (b) When a sentence conforms to a template but for one contradicting piece of information, an attempt is made to make the sentence fall in line with the template, by reinterpreting another one of its elements.

We believe that we have offered the first demonstration of the power of utility templates in the interpretation of conditional sentences. In the rest of this article, we discuss the relation between our data and other content-specific approaches to conditionals. In a first section, we compare our approach to that of pragmatic reasoning schemas, Darwinian algorithms, and decision-theoretic approaches to conditional reasoning. In a second section, we relate utility templates to Relevance Theory, by way of addressing the question of how to generate new potential templates in a principled way. Specifically, we introduce the folk axioms of decision that allow for the derivation of inferences from utility grids, and we show that the templates we

identified share one characteristic: They maximize the inferential potential of the conditional statement with respect to these folk axioms.

*Content-specific approaches to conditionals*

Content-specific approaches to conditionals have long recognized the importance of preferences and utilities. Pragmatic reasoning schemas for permission and obligation (Cheng & Holyoak, 1985), Darwinian algorithms for cheater detection (Cosmides et al., 2010), decision-theoretic models of hazard management (Perham & Oaksford, 2005) or slippery slope argumentation (Corner et al., 2011), all apply to conditionals whose clauses feature preferred or non-preferred propositions.

The theory of utility conditionals (Bonnefon, 2009) differed from previous approaches in the breadth of its coverage: It offered the possibility to classify and formalize all utility conditionals, rather than focusing on a specific subset of situations (e.g., promises, precautions). In a sense, the identification of utility templates brings the theory closer to its content-specific predecessors, by re-focusing attention to specific subsets of utility conditionals such as Social Contracts. After the expansion phase that corresponded to the theory of utility conditionals (which identified hundreds of utility patterns besides that studied in previous research), the discovery of utility templates corresponds to a contraction phase (by showing that large subsets of patterns can be subsumed under a tractable number of templates).

One question we have not addressed thus far is how utility templates emerge. In this respect, our take on utility template differs from that of previous content-specific approaches. These previous approaches were divided about the nature (innate or learned) of the cognitive mechanisms elicited by specific, utility-relevant situations such as cheater detection or hazard management. We remain agnostic about this issue, as far as utility templates are concerned. We believe that the emergence of utility templates might result from distinct mechanisms (Darwinian importance, mere frequency, pragmatic relevance – see next section), and that the

emergence of a given template might quite possibly be multi-determined. In fact, multiple determination could explain why some templates may exert greater constraints than others.

We have to remain cautious about this last phenomenon. Although we consistently found that our experimental effects were greater for some templates than for others, the exact interaction effects were not always consistent from one study to another. For example, Social Contracts delivered comparatively strong effects in Study 1, moderate to strong effects in Study 2, and moderate to comparatively weaker effects in Study 3. The Unpacking template was the only one to deliver comparatively stronger effects across all three studies. Future research might be able to offer an explanation for this finding – but this task would be made easier if we could come with a complete list of utility templates. Only when we can be reasonably sure that we have identified most utility templates, will we be on solid grounds to identify the characteristics of these templates which exert the strongest constraints on interpretation. This begs the question of how to generate a list of possible templates, in a principled way, rather than in the exploratory way we adopted in our completion survey. We address this issue in the next section.

#### *Utility templates maximize inferential potential*

Let us first detail how utility grids invite conclusions. Utility grids capture patterns of agents, actions and valued outcomes. In the theory of utility conditionals (Bonnenfon, 2009), these grids invite conclusions by means of *folk axioms of decision*, that is, naive beliefs about the way agents make their decisions. Bonnenfon (2009) introduced three such folk axioms: self-interested behavior, self-interested attitude, and limited altruism.

Table 4 displays these three folk axioms, together with the grids that trigger them, and the conclusion invited from each triggering grid. When a grid matches one of the 12 patterns in Table 4, it invites the conclusion shown immediately under the grid. A single grid can match several of the patterns in Table 4. In that case, inconsistent conclusions cancel out, and all

remaining conclusions are invited; furthermore, conclusions are stronger if they are invited through several matches.

Let us now define the inferential potential of a grid, or more precisely the order relation expressing that a grid has greater inferential potential than another. Let us call  $\mathcal{C}_1$  the set of conclusions invited by folk axioms from Grid 1, and  $\mathcal{C}_2$  the set of conclusions invited by folk axioms from Grid 2. We will say that Grid 1 has greater inferential potential than Grid 2 when it invites a greater number of conclusions ( $|\mathcal{C}_1| > |\mathcal{C}_2|$ ), or when the two grids invite the same number of conclusions and there is a conclusion in  $\mathcal{C}_1$  that is stronger than any conclusion in  $\mathcal{C}_2$ .

Equipped with this definition, we can identify a characteristic shared by the four utility templates we have considered so far: Template-congruent interpretations in Studies 1 and 2 always had greater inferential potential than their template-incongruent counterparts. We consider in detail the case of the Unpacking template, and let the reader check that the other templates behave in similar fashion. Let us consider a statement that could fit the Unpacking template:

(15) If Alan peens Claire, then she will be dested. Claire dislikes being peened.

A template-congruent interpretation of (15) would be that Claire dislikes being dested:

$$\left\{ \begin{array}{l} a - c \\ \omega - c \end{array} \right\},$$

and a template-incongruent interpretation of (15) would be that Claire likes being dested:

$$\left\{ \begin{array}{l} a - c \\ \omega + c \end{array} \right\}.$$

The template-congruent grid does not match any of the grids associated with the folk

axiom of self-interested behavior, but it matches two of the grids associated with the folk axiom of self-interested attitude (third and fourth grids, both invites the conclusion that Claire thinks Alan should not peen her), as well as two of the grids associated with the folk axiom of limited altruism (third and fourth grids, both invites the conclusion that Alan will not peen Claire).

Thus, the statement strongly invites two conclusions: Claire thinks Alan should not peen her, and Alan will not peen Claire.

In contrast, the template-incongruent grid does not invite any conclusion based on folk axioms. First, it does not match any of the grids associated with self-interested behavior. Second, while it matches two of the grids associated with self-interested attitude (second and third grids), these grids support inconsistent conclusions (Claire thinks Alan should peen her, and Claire thinks that Alan should not peen her). These conclusions cancel out. Third, while the template-incongruent grid matches two of the grids associated with limited altruism (second and third grids), it is again the case that these grids support inconsistent conclusions that cancel out. In sum, the template-incongruent grid does not support any conclusion.

A similar analysis applies to the three other templates (for the sake of brevity, we will not provide this full analysis). In each case, the template-congruent has greater inferential potential than the template-incongruent interpretation. This result offers two interesting perspectives: A principled way to identify potential templates, and a direct link to a relevance-theoretic approach of utility templates. We consider these two perspectives in the next section.

### *Perspectives*

Our exploratory approach allowed us to identify four utility templates, but it did not provide us with a principled way to generate new candidates. The observation that our four templates appeared to be such that they maximized inferential potential provides us with such a principled approach: To identify new utility templates, we should search within the set of

utility grids with the highest inferential potential. This would require to rank-order all well-formed utility grids by inferential potential, and to work our way from the top. Although such an enterprise is beyond the scope of the current discussion, we can consider as an illustration a grid that is among the strongest in terms of inferential potential, which we may call the Equal Treatment grid:

$$\left\{ \begin{array}{l} x + y \\ x + z \end{array} \right\}.$$

This grid reflects statement such as:

- (16) a. If Alice helps Bob, then she will help Charles;  
 b. If I give you a bonus, then I will give Tom a promotion.

This grid matches two of the grids associated with self-interested attitude, and two of the grids associated with limited altruism. For example, (16-a) would thus invite the conclusion that Bob thinks Alice should help him, invite the conclusion that Charles thinks Alice should help Bob, and strongly invite the conclusion that Alice will help Bob. Having identified the Equal Treatment grid as a potential utility template, the next step would be to check whether the grid produces the two characteristic effects of utility templates, as observed in Studies 1 and 2.

Beyond its role in identifying new templates, the hypothesis that utility templates maximize inferential potential provides an interesting connection with Relevance Theory (Sperber & Wilson, 1986/1995). Relevance Theory postulates that statements convey an assumption of optimal relevance and are interpreted as such, that is, statements receive the interpretation that maximize their relevance to the listener. All other things being equal, greater processing effort decrease the relevance of an interpretation, and greater cognitive effects increase the relevance of an interpretation.

Remarkably, cognitive effects can be broadly construed as inferential potential. That is,

the cognitive effects of an interpretation are greater when this interpretation allows one to reach a greater number of conclusions, or to reach stronger conclusions. Our current hypothesis (utility templates are found among grids that maximize the inferential potential of utility conditionals) might thus be rephrased as: Utility templates are found among grids that maximize the relevance of utility conditionals. The two characteristic effects of utility templates could then be reduced to a relevance effect: Utility templates constrain interpretation and revision by virtue of being more relevant than their close variants.

Note though that this perspective remains speculative at this stage. Importantly, the inferential potential of a utility grid (and thus, the relevance of the corresponding interpretation of the conditional) depends on the number of folk axioms of decision that apply to the grid. The list of folk axioms of decision, however, is not closed, and may be modified by future research. An investigation of the relevance account of our findings is thus highly dependent on the progress of research regarding the folk theory that people entertain about the way other agents make their decisions.

### *Conclusion*

Utility conditionals describe configurations of agents, actions and valued consequences. Utility templates are a special subset of these configurations, which exert strong constraints on interpretation. When a situation is close enough to a template, people interpret ambiguous information or reinterpret current information in such a way that their understanding of the situation fits the template. We have identified a set of four templates, and offered the first empirical demonstration of their characteristic effects. We have considered a process explanation of these effects, which allows for the principled generation of other templates, but also for a possible reformulation of our findings within the framework of relevance theory.

## References

- 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.
- Bonnefon, J. F. (2009). A theory of utility conditionals: Paralogical reasoning from decision-theoretic leakage. *Psychological Review*, 116, 888–907.
- Bonnefon, J. F. (2012). Utility conditionals as consequential arguments: A random sampling experiment. *Thinking and Reasoning*, 18, 379–393.
- Bonnefon, J. F., Girotto, V., & Legrenzi, P. (2012). The psychology of reasoning about preferences and unsequential 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.
- Bonnefon, J. F., & Sloman, S. A. (in press). The causal structure of utility conditionals. *Cognitive Science*.
- Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, 17, 391–416.
- Corner, A., Hahn, U., & Oaksford, M. (2011). The psychological mechanism of the slippery slope argument. *Journal of Memory and Language*, 64, 153-170.
- Cosmides, L., Barrett, H. C., & Tooby, J. (2010). Adaptive specializations, social exchange, and the evolution of human intelligence. *Proceedings of the National Academy of Science of the USA*, 107, 9007–9014.
- Evans, J. S. B. T. (2005). The social and communicative function of conditional statements. *Mind & Society*, 4, 97–113.
- Evans, J. S. B. T., Neilens, H., Handley, S. J., & Over, D. E. (2008). When can we say 'if'? *Cognition*, 108, 100–116.
- 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.
- Hilton, D. J., Kemmelmeier, M., & Bonnefon, J. F. (2005). Putting ifs to work: Goal-based relevance in conditional directives. *Journal of Experimental Psychology: General*, 135, 388–405.
- Legrenzi, P., G., & Girotto, V. (1996). Contract proposals: a sketch of a grammar. *Theory & Psychology*, 6, 247–265.
- López-Rousseau, A., Diesendruck, G., & Benozio, A. (2011). My kingdom for a horse: On incredible promises and unpersuasive warnings. *Pragmatics and Cognition*, 19, 399–421.
- 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.
- Ohm, E., & Thompson, V. (2004). Everyday reasoning with inducements and advice. *Thinking and Reasoning*, 10, 241–272.
- Perham, N. R., & Oaksford, M. (2005). Deontic reasoning with emotional content: Evolutionary psychology or decision theory? *Cognitive Science*, 29, 681–718.
- Politzer, G., & Nguyen-Xuan, A. (1992). Reasoning about conditional promises and warnings: Darwinian algorithms, mental models, relevance judgements or pragmatic schemas? *Quarterly Journal of Experimental Psychology*, 44, 401–412.
- Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC Nonword Database. *Quarterly Journal of Experimental Psychology*, 55A, 1339–1362.
- Sperber, D., & Wilson, D. (1995). *Relevance, communication and cognition*. Oxford: Blackwell. (Original work published 1986)
- Thompson, V. A., Evans, J. S. B. T., & Handley, S. J. (2005). Persuading and dissuading by conditional argument. *Journal of Memory and Language*, 53, 238–257.

## Appendix

### Conditional fragments used in the completion survey

1. If I give you a job, then I . . .
2. If you give me a job, then I . . .
3. If I give you a job, then you . . .
4. If you give me a job, then you . . .
5. If I hurt you, then I . . .
6. If you hurt me, then I . . .
7. If I hurt you, then you . . .
8. If you hurt me, then you . . .
9. If you buy yourself a new car, then you . . .
10. If you buy yourself a new car, then I . . .
11. If I buy myself a new car, then you . . .
12. If I buy myself a new car, then I . . .
13. If you self harm, then you . . .
14. If you self harm, then I . . .
15. If I self harm, then you . . .
16. If I self harm, then I . . .
17. If you stand next to me, then I . . .
18. If you stand next to me, then you . . .
19. If you move two inches, then you . . .
20. If you move two inches, then I . . .
21. If I stand next to you, then I . . .
22. If I stand next to you, then you . . .
23. If I move two inches, then I . . .

24. If I move two inches, then you...
25. If John helps Sarah, then he...
26. If Colin helps Laura, then she...
27. If Brian insults Mandy, then he...
28. If Robert insults Joanne, then she...
29. If Chris talks to Alison, then he...
30. If Paul talks to Leanne, then she...

### **Author Note**

Address correspondence to Jean-François Bonnefon, 'Cognition, Langues, Langage et Ergonomie,' Maison de la recherche, 5 allées A. Machado, 31058 Toulouse Cedex 9, France. E-mail: [bonnefon@univ-tlse2.fr](mailto:bonnefon@univ-tlse2.fr). This research was partially supported by a Leverhulme Trust grant (F/00 120/BT) awarded to the third author.

### Footnotes

<sup>1</sup>This convention allows one to give a utility grid to conditionals such as 'If the figure is a triangle, then the sum of its interior angles is 180 degrees'. In this example,  $p$  and  $q$  are simply noted as having zero utility to the whole set of agents. For such and other issues, such as situations in which a single action has different degrees of utility to different agents, see Bonnefon (2009).

Table 1

*Utility templates suggested by the sentence completion survey. In all grids,  $u$  is different from zero.*

Justice	Social Contract	Unpacking	Tautology
13% of completions	11% of completions	10% of completions	2% of completions
$\begin{pmatrix} x & u & y \\ \omega & u & x \end{pmatrix}$	$\begin{pmatrix} x & u & y \\ y & u & x \end{pmatrix}$	$\begin{pmatrix} x & u & y \\ \omega & u & y \end{pmatrix}$	$\begin{pmatrix} x & u & y \\ x & u & y \end{pmatrix}$

Table 2

*Study 1: Mean target nonverb ratings as a function of template and valence of the companion nonverb. Standard deviation is shown within parentheses.*

Template	Companion nonverb valence	
	Positive	Negative
Unpacking	+1.8 (1.2)	-1.7 (1.3)
Social Contract	+1.6 (1.1)	-1.5 (1.3)
Tautology	+1.5 (1.4)	-1.3 (1.4)
Justice	+1.3 (1.3)	-1.4 (1.1)

Table 3

*Study 4: Perceived utility of the target verbs ('Praise' and 'Hurt'), as a function of the connective featured in the sentence, and the congruency of the companion nonverb.*

	Praise		Hurt	
	Control	Conflict	Control	Conflict
Conjunctions	+1.2 (1.4)	+1.2 (1.5)	-1.4 (1.5)	-1.0 (1.6)
Disjunctions	+1.2 (1.5)	+0.9 (1.6)	-1.1 (1.6)	-0.8 (1.8)

Table 4

Three folk axioms of decision, together with the utility grids that trigger them and the conclusions they invite whence triggered by these grids. The black dot is a blank parameter that can take any legitimate value.

**Self-interested behavior.** People take actions that increase their own personal utility, and they do not take actions that decrease their own personal utility.

$$\begin{Bmatrix} x & + & x \\ \bullet & \bullet & \bullet \end{Bmatrix}$$

$x$  will do  $p$

$$\begin{Bmatrix} x & \bullet & \bullet \\ \bullet & + & x \end{Bmatrix}$$

$x$  will do  $p$

$$\begin{Bmatrix} x & - & x \\ \bullet & \bullet & \bullet \end{Bmatrix}$$

$x$  will not do  $p$

$$\begin{Bmatrix} x & \bullet & \bullet \\ \bullet & - & x \end{Bmatrix}$$

$x$  will not do  $p$

**Self-interested attitude.** People think that actions that increase their own personal utility should be taken by others, when these other agents can take these actions; and they think that actions that decrease their own personal utility should not be taken by others, when these other agents can take these actions.

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

$y$  thinks  $x$  should do  $p$

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

$y$  thinks  $x$  should do  $p$

$$\begin{Bmatrix} x & - & y \\ \bullet & \bullet & \bullet \end{Bmatrix}$$

$y$  thinks  $x$  should not do  $p$

$$\begin{Bmatrix} x & \bullet & \bullet \\ \bullet & - & y \end{Bmatrix}$$

$y$  thinks  $x$  should not do  $p$

**Limited altruism.** People take actions that increase the utility of others, insofar as doing so does not decrease their own personal utility; and they do not take actions that decrease the utility of others, insofar as these actions do not increase their own personal utility.

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

$x$  will do  $p$

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

$x$  will do  $p$

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

$x$  will not do  $p$

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

$x$  will not do  $p$

where  $(u, z) \neq (-, x)$

where  $(u, z) \neq (+, x)$

### Figure Captions

*Figure 1.* Study 2: Perceived utility of the target verbs ('Help' and 'Hurt'), as a function of the utility template and congruency of the companion nonverb. Error bars show the standard error of the mean.

*Figure 2.* Study 3: Perceived utility of the target verbs ('Help' and 'Hurt'), as a function of the utility template and congruency of the companion nonverb. Error bars show the standard error of the mean.



