Moral Structure Falls Out of General Event Structure

Brent Strickland, Matthew Fisher & Joshua Knobe

Yale University

> *Note*: This paper was originally published as a response to a target article by Kurt Gray, Liane Young and Adam Waytz. However, we have tried our best to write the paper in such a way that it will be easily accessible to anyone with an interest in these issues (even those who have not yet read the original target article).

At the core of Gray, Young and Waytz's fascinating and wide-ranging target article is the idea that people understand morally charged events in terms of two basic roles. On one hand, there is an *agent* who intentionally brings about an outcome; on the other, there is a *patient* who suffers that outcome's effects.

Much of the article is devoted to working out the surprising empirical implications of these role assignments, exploring their relevance to everything from psychopathology to religious belief. This work raises many deeply interesting issues, but we will be focusing here on a question that is in some sense more fundamental. We want to know why it is that people understand moral events in terms of these specific roles in the first place. Why is it that people consistently make sense of moral events in terms of agents and patients instead of thinking in terms of some other set of roles, or perhaps even using different roles to understand different events?

The best way to convey the force of this question might be to introduce a sports analogy. If we look at a typical sport, we find that each player is assigned to a specific role,

with its own unique responsibilities and distinctive prohibitions. But of course, the roles vary radically from one sport to the next. We find one set of roles in baseball (pitcher, catcher, shortstop), another in football (quarterback, running back, linebacker), and yet a third in soccer (goalie, midfielder, forward).

But imagine for a moment that things had turned out differently. Suppose it had turned out that every sport, in every culture throughout the world, made use of the same basic set of roles. Suppose that every single sport assigned the players to be pitchers and catchers, and the only difference between sports lay in the precise details of what the players in these two roles were supposed to do. This would be a deeply surprising outcome, and it would leave us with a difficult theoretical question. We would want to know why the different sports all ended up converging on this same underlying structure.

What Gray and colleagues are suggesting in the domain of morality is something more or less along the lines of what we just imagined in the domain of sports. There is an enormous variety of different acts that people regard as morally wrong – everything from murder to dishonesty to incest – but the theory is that all of these acts are understood in terms of the same basic set of roles. No matter which immoral act you consider, there will always be an agent and a patient. It therefore becomes possible, they suggest, to develop a very general understanding of agency and patiency that can be applied across all possible immoral acts.

Gray and colleagues make a strong case for their central claim, but they thereby leave us immediately with a new question. Why should it be that all of these different transgressions are understood in terms of the same basic roles? Are these roles specific to morality, or do they result from more general mental processes that happen to encompass the understanding of moral events?

Returning to our analogy, if we found a commonality among all sports, it seems that we could explain this commonality in either of two ways. One approach would be to focus specifically on sports. For example, one could hypothesize that humanity designed all sports in this particular way so that they would fulfill a function that was specifically related to sports events (e.g. pleasing audiences). By contrast, a second approach would be to suggest that the patterns observed in sports can be subsumed under a much broader regularity. For example, one could suggest that the features that at first seemed to be specific to sports actually arose for most organized human activities. So the answer to the mystery of why sports work the way they do would ultimately depend on what we learned about organized human activities more generally.

Similarly, it seems that there are two possible ways of making sense of commonalities in the way people think about morally charged events. One approach would be to focus on moral judgments in particular. We might start with the observations Gray and colleagues make about the nature of moral cognition, and ask how the details of moral role assignment differ according to attributes like age, gender or hero/villain-hood status. Ultimately, we might be able to explain the existence of agent and patient roles in terms of the specific features of cognition within the moral domain.

This strategy is certainly a plausible one, and has proved valuable in discovering new psychological findings. Here, however, we will be arguing for a very different approach. We suggest that the role assignments one finds in moral cognition can be explained in terms of a far more general theory about how people make sense of events. The claim will be that people show a quite general tendency to construe events in terms

of agents and patients.[1] The effects we observe in the moral domain then fall out of that general tendency.

To illustrate this claim, let's consider a simple example. Suppose you enter a room and see that John is selling crack cocaine to Stephen. On the theory developed by Gray and colleagues, your evaluation of this act will depend in large part on whether you construe Stephen as an agent or a patient. One way to understand the situation would be to see Stephen as a person who willfully and intentionally decided to buy illegal drugs; another would be to see Stephen as the hapless victim of a drug dealer who is being taken advantage of. You will arrive at radically different moral judgments depending on which of these construals you adopt.

Our claim now is that this distinction between agents and patients actually arises from a much more general fact about human cognition. Suppose you came into the room and simply found John selling Stephen a cucumber. In that case, there would be no moral issues to confront – no question about right or wrong, praise or blame. Nonetheless, we suggest that you would still be able to construe Stephen either as an agent or as a patient and that this construal would involve the same processes we see at work in more morally charged cases. In other words, the hypothesis is that the structure one finds in moral cognition actually falls out of a structure one finds in event cognition more generally.

To test this hypothesis, we begin by reviewing research on the assignment of roles in events generally and then conduct a study that explores the link between general event cognition and moral cognition.

---

[1] It is however important to point out that it is of course possible to also construe events as having just agents (as in "John ran"), just patients (as in "The window was hit"), or as having additional roles such as instruments or locations. Our point here is just that the agent/patient dyad is a very common way to construal of events.

Syntax designates agents and patients

To reach a better understanding of the assignment of roles in general event cognition, we will be turning to a perhaps unexpected source. Specifically, we will be looking to the literature in *theoretical linguistics*.[2] The hope is that facts about how people use language can offer us a window into more general questions about how people understand agency and patiency.

To begin with, consider the following three sentences:

(1) The dog bit the man.

(2) The foreman ate his lunch.

(3) The woman smashed the window.

These three sentences describe three very different kinds of events, but they nonetheless display a similar structure. In each of the sentences, one thing appears as the *subject* (the dog, the foreman, the woman) and another appears as the *object* (the man, his lunch, the window). The claim now will be that the structure of these sentences can give us an important clue as to how people represent events.

The first thing to notice is that it does not appear to be entirely arbitrary which thing is assigned to be the subject and which the object. Instead, there seems to be a strikingly consistent pattern. Though the three sentences describe three very different

[2] The literature on these topics in linguistics can be quite technical and complex, and we therefore introduce a few simplifications to aid exposition. First, within the literature on thematic roles, there are a number of competing theories (Dowty, 1991; Carlson & Tanenhaus, 1988; Fillmore, 1968), but since the differences between those theories are not relevant to the specific issues discussed here, we will be ignoring them. Second, we will generally be less than explicit about the distinction between objects in the world and the linguistic expressions that denote those objects. (So instead of writing 'a linguistic expression that denotes John appears as the grammatical subject,' we will use the shorthand 'John appears as the grammatical subject.') Finally, it is conventional to write the names of thematic roles in all caps (e.g., AGENT), but since much of our paper is specifically concerned with the question as to whether thematic roles differ from roles of other types, we will be flouting that convention here.

kinds of events, we always find that the subject shows the same basic properties. In all three cases, the subject is the one who initiated the action, who caused the outcome, and who acted more intentionally.

One might assume at first that this pattern just happens to come out in our three examples and that there could just as easily be sentences that work the opposite way. For example, suppose we define a verb *shmite* whose meaning is simply the converse of *bite* (so that it means something like *be bitten by*). Then it might seem that we could switch everything around and use a sentence like:

(1)′ The man shmit the dog.

[which would have the same meaning as 'The dog bit the man.']

Yet research in linguistics suggests otherwise. Sentences like this one do not appear to occur in any human language. In other words, it seems that one can't just arbitrarily assign any old thing to be the subject of the sentence when both an agent and patient are explicitly mentioned. Instead, there appear to be general principles that tell us which things will appear as subjects and which as objects (Baker, 1997). As long as people understand these general principles, they can see immediately that it is acceptable to have a verb like *bite* but that there is something wrong with a verb like *shmite*.

If we now want to characterize these general principles, we will have to construct a theory at a more abstract level. It won't be enough just to talk about each verb separately. (We can't just have one rule for the verb *bite*, another for the verb *ate*, and so on.) What we need is rather a very general theory that will apply across all of these different verbs. The theory should tell us, at a very general level, how to figure out which things will appear as subjects and which as objects.

To address this question, linguists have posited a system of abstract 'thematic roles' (Dowty, 1991; Carlson & Tanenhaus, 1988; Fillmore, 1968). Such roles include *theme*, *experiencer*, *instrument* and a number of others, but for present purposes, we will be especially concerned with the role of *agent*. Roughly speaking, something counts as an agent to the extent that it causes events to happen, intends them to happen, and initiates their occurrence (Dowty, 1991). When people classify something as the agent in a given sentence, they generally include it as the subject of that sentence.

With this conceptual framework in hand, it becomes possible to offer a new account of how people use those sentences. Consider again the sentence:

(1) The dog bit the man.

To use this sentence correctly, people do not need to learn a special rule about which thing counts as the subject in sentences involving biting. Instead, all they need is a very abstract understanding of how to determine which thing counts as the agent of an event. They can see that the dog is the agent of the biting, and they therefore know that it has to appear as the subject of the sentence.

In sum, research in linguistics suggests that people show a general tendency to understand events in terms of agents and patients. This tendency is not limited to cases that have moral significance. On the contrary, it appears to be a pervasive feature of human cognition – one that can be found not only when we are making moral judgments but also when we are simply trying to understand ordinary events.

Thematic roles and moral roles

Thus far, we have been examining two different kinds of theories that involve role assignments. Gray and colleagues have developed a theory about role assignment in moral psychology, while linguists have developed theories about role assignment in the

understanding of events more generally. We now want to ask about the relationship between these two kinds of theories. Are the roles we find in moral psychology the very same roles that can be found in people's ordinary understanding of events, or are they of a categorically different class?

To address this question, we need to form a bridge between the methods of linguistics and the topics of moral psychology. That is, we need to manipulate role assignments using the usual linguistic methods (subject vs. object of a sentence) and then check to see whether this manipulation has any impact on people's intuitions about the usual moral questions (intentionality, suffering, punishment). Ideally, we would find that whichever person is assigned to be the subject of a sentence is regarded as having all of the trademark features that Gray and colleagues associate with moral agency – more intentionality, less suffering, more responsibility – regardless of the moral status of the event itself.

It may seem at first, however, that such a study is impossible to arrange. After all, suppose we simply looked at the following pair of sentences:

(4) The dog bit the man.

(5) The man bit the dog.

It is true that these sentences differ in the linguistic position given to the man (he is the object in one, the subject in the other). But these sentences also differ in some much more salient respects: they describe two radically different kinds of events, and they will therefore be seen as differing, for obvious reasons, in their moral status. What we really want, then, is a pair of sentences that are more tightly controlled. We want a pair of sentences that allow us to vary the linguistic features without also changing the nature of the event described.

Luckily, however, it is possible to construct pairs of sentences that control for these other differences. Consider the situation(s) described by sentences (6) and (7) below:

(6) John sold products to Susan.

(7) Susan bought products from John.

These two sentences are logically symmetric and thus describe the same set of possible circumstances. There is a clear sense in which the two sentences are saying exactly the same thing about which event occurred. Yet, all the same, there is an important difference between the two. The first sentence assigns John to the role of agent while the second gives this role to Susan.

Recent work in our lab has revealed that when people actually interpret these sentences they have an automatic bias to attribute more agency to grammatical subjects compared to non-subjects (Strickland, Fisher, Keil & Knobe, 2012; Strickland, Fisher, Peyroux & Keil, 2011). So for example, in one experiment, participants were shown sentences like (6) or (7) and asked to indicate whether or not the underlined person acted intentionally or unintentionally. They either made these judgments under time pressure or with encouragement to take their time and think deeply about each sentence. When probed about grammatical subjects, participants showed a significantly stronger bias to indicate that the person acted more intentionally in the speeded compared to the unspeeded condition. For grammatical objects (like "John" in sentence (7)) however, no such bias existed. This finding thus suggests an asymmetry in the automatic attribution of intentionality, with syntactic agents receiving more of it than syntactic patients.

While being logically symmetric, verb pairings like "buy" and "sell" are *lexically* asymmetric in the sense that a change in verb is necessary to get a logical entailment like the one present in sentences (6) and (7) (whereby the logic of both sentences is the same

but "John" and "Susan" switch grammatical roles). However, similar effects obtain in perfectly symmetric verbs like "marry", "french kiss", "trade", "swap", "date" and "make love".

Each of these verbs is logically symmetric because both people are doing the same action. So consider the morally neutral scenario described by sentence (8) below:

(8) John and Susan french kissed.

John and Susan are both agents, but you can also describe this scenario by either sentence (9) or (10) below:

(9) John french kissed Susan

(10) Susan french kissed John.

Given the nature of french kissing, it seems that these two sentences are completely equivalent. (There is no way for it to be the case that John french kissed Susan unless it is also the case that Susan french kissed John.) Yet, once again, there is a difference in the assignment of thematic roles. John appears as an agent in the first sentence but not in the second, and one might expect that this difference would affect people's understanding of the event described.

Indeed this was confirmed in our studies (Strickland, Fisher, Keil & Knobe, 2012; Strickland, Fisher, Peyroux & Keil, 2011). So, for example, in a simple judgment task when participants were shown just one sentence and asked to indicate how intentionally either the grammatical subject or direct object acted, they consistently judged subjects to be more intentional than objects. But when asked about either actor in a sentence like (8), they rated both as acting equally intentionally.

What we have here, then, is a method that makes it possible to hold almost everything constant while varying only the assignment of thematic roles. A reader who is

confronted with sentence (9) has been told about almost exactly the same event as a reader confronted with sentence (10). The primary difference between the two is just that the two sentences differ in which person has been assigned to be the agent.

It therefore becomes possible to put our hypothesis to the test. The trick is to construct pairs of sentences that are similar to the ones we have already examined except that they appear both in the morally charged and morally neutral domains. Then we can use those sentences to examine the relationship between thematic roles and moral roles. The basic technique would be to manipulate the thematic roles using the usual linguistic method (subjects vs. objects) and then explore the impact on moral roles using the dependent variables that Gray and colleagues have developed in their earlier work (intentionality, suffering and punishment).

Experiment

In order to test the influence of syntax on these mental state attributions, we designed a simple experiment in which we probed participants' intuitions about the mental states of characters from a described event. We manipulated the grammatical position (subject vs. direct object) of the person being asked about, and did this in both morally charged and morally neutral scenarios.

In addition to asking about traits associated with agency (intentionality and responsibility), we also asked about a trait associated with being a patient (being upset). In the same way that the association between grammatical subjects and the role of agent could lead to increased perceptions of intentionality and responsibility for subjects, perhaps the association between the direct object position and the role of patient could lead to increased perceptions of suffering for objects, even in morally neutral scenarios.

If these general syntactic cues to agency and patiency went on to influence moral

judgments, this would be even stronger evidence for the view that moral dyads are grafted onto a more general event structure. For example, it could be the case that grammatical subjects are generally seen as acting more intentionally than grammatical objects, and moral judgments like punishment are sensitive to this more basic asymmetry.

There were a few important constraints on our experimental design. First, it was necessary to create matched pairs in which would allow us to manipulate the variables of interest in highly similar morally charged and morally neutral scenarios. Secondly, it was necessary that the action described be logically symmetric such that the grammatical subject and object could be reversed (i.e. with the object becoming the subject and vice-versa) without changing the true meaning of the sentence. In order to do this, we employed the verb "french kiss" which has the necessary logical symmetry. To create our moral vs. morally neutral contrast, we described the two participants as either being a 25 year old male and 15 year old female (who would thus be breaking many U.S. state laws) or as a 35 year old male and 25 year old female (who would not theoretically break any state laws).

So participants rated dimensions associated with agency, patiency and blame-worthiness for either the grammatical subject or object who was either an older male or a younger female. Each participant did this for both morally charged ((5) and (6)) and for morally neutral ((6) and (7)) scenarios.

<u>MORALLY CHARGED</u>

(1)   Steven is 25 years old and Kate is 15 years old.

        Steven french kissed Kate.

(2)   Steven is 25 years old and Kate is 15 years old.

        Kate french kissed Steven.

MORALLY NEUTRAL

(3)   Steven is 35 years old and Kate is 25 years old.

        Steven french kissed Kate.

(4)   Steven is 35 years old and Kate is 25 years old.

        Kate french kissed Steven.

It follows naturally from our experimental design that in addition to manipulating syntactic cues to agency and patiency, we also end up manipulating social cues like age and gender that could potentially trigger the assignment of specifically moral roles. For example, previous research has shown that the young people and women are more likely to be viewed as victims of crimes, and mature adults and men are more likely to be viewed as perpetrators (Gray and Wegner, 2009; Howard, 1984). Thus if there exist social cues to agency/patient-hood in specifically moral scenarios, then age and gender would be likely candidates.

Methods

Participants

83 paid online participants from Amazon's Mechanical Turk.

Materials and Procedure

This study used a 2 (syntactic position) x 2 (social cue) x 2 (morality) design with syntactic position and social cue as between subject factors and morality as a within subjects factor.

Each participant made judgments about two passages describing an older male and younger female french kissing (see sentences 1-4 above). The first sentence of the passage introduced the characters and their social characteristics. The order of mention

of the two characters in the introductory sentence was randomly assigned for each

passage. The order or presentation of the moral and morally neutral passages was

counterbalanced between participants.

For each of the two scenarios, participants rated the relevant character on four

variables (always using a 7-point scale) with the following questions:

How intentionally did Steven act?
How responsible was Steven for the act?
How likely is Steven to be upset?
How harshly should Steven be punished?

The order of these questions was counterbalanced.

Results

Means and standard deviations for all dependent variables are reported in Table

1. Analyses were conducted separately for each dependent measure. For each of these

measures, participants' ratings were submitted to 2 (syntactic position) x 2 (social cue) x 2

(morality) mixed-model ANOVA, with syntactic position and social cue as between-

subjects factors and morality as a within-subjects factor.

For each dependent measure, the primary questions of interest were (a) whether

there would be an effect of syntactic position and (b) whether that effect would arise both

for morally charged and morally neutral scenarios. We found no significant effects of

either participant gender or the syntactic position of the queried actor in the introductory

statement, and those variables will not be discussed further.

Table 1: Mean Responses (Standard Deviations) to Dependent Measures

	Intentionality	Responsibility	Upset	Punishment
Morally Charged				
Older/Male				
Subject	6.11 (1.52)	5.42 (2.04)	2.32 (1.49)	4.63 (1.92)
Object	3.41 (1.92	3.68 (2.03)	4.27 (1.60)	2.82 (1.97)
Younger/Female,				
Subject	5.86 (1.55)	5.27 (1.28)	3.36 (2.22)	3.95 (2.15)
Object	3.30 (1.42)	3.00 (1.49)	3.80 (1.70)	2.30 (1.42)
Morally Neutral				
Older/Male				
Subject	5.53 (1.50)	5.53 (1.61)	1.74 (1.19)	1.74 (1.19)
Object	4.23 (1.74)	3.59 (2.06)	2.23 (1.60)	1.23 (.61)
Younger/Female,				
Subject	6.36 (1.40)	6.27 (1.20)	1.64 (1.09)	1.27 (.63)
Object	4.35 (1.76)	4.40 (1.70)	2.00 (1.12)	1.05 (.22)

Intentionality

There was a main effect of syntactic position such that subjects of sentences were seen as acting more intentionally than objects, $F(1,79)=45.59$, $p<.001$. There was also a main effect for morality such that actors in the morally neutral scenario received higher intentionality ratings than actors in the morally charged scenario, $F(1,79)=7.86$, $p=.006$. There was no significant effect of social cue. There was a significant interaction of morality and social cue, $F(1,79)=4.22$, $p<.05$, such that the social cues only had an effect in the morally neutral scenario. In fact, it was the younger female who was rated as more intentional in the morally neutral charged scenario. Finally, there was a significant interaction between morality and syntactic position, $F(1,79)=9.30$, $p=.003$. No other interactions were significant.

An inspection of the means showed that the morality x syntactic position interaction arose because the impact of syntactic position was greater for the morally charged scenario than for the morally neutral scenario. To further examine this interaction pattern, we looked separately at the two scenarios. Since there was no significant effect of social cue, we collapsed across this variable and conducted separate t-tests. The results showed a significant effect of syntactic position both for the morally charged scenario, t(81)=7.43, p<.001, and for the morally neutral scenario, t(81)= 4.76, p<.001.

Responsibility

Participants rated syntactic subjects as more responsible than syntactic objects, F(1,79)=33.21, p<.001. There was also a main effect of morality such that actors in the morally neutral charged scenario received higher responsibility ratings than actors in the morally neutral scenario, F(1,79)=14.30, p<.001. There was no main effect of social cue, but there was a significant morality x social cue interaction such that social cue had a greater influence on responsibility ratings in the morally neutral scenario, F(1,79)=13.96, p<.001. There was no interaction of syntactic position and morality.

When the two scenarios were examined separately, grammatical subjects were rated as more responsible than objects both for the morally charged, t(81)=5.22, p<.001, and for the morally neutral scenario, t(81)= 5.23, p<.001.

Upsetness

Grammatical objects were seen as more upset than grammatical subjects, F(1,79)=7.75, p<.01. There was also a main effect of morality such that participants rated the total level of upsetness to be higher in the morally charged scenario than in the morally neutral scenario, F(1,79)=74.57, p<.001. There was no main effect of social cue

or any significant interaction between social cue and the other variables. There was a significant morality x syntax interaction, $F(1,79)=4.67$, p<.05. No other interactions were significant.

An inspection of the means showed that the impact of syntactic position on upsetness ratings was greater in the morally charged scenario than in the morally neutral scenario. Looking separately at the two scenarios, we found that the effect of syntactic position was significant in the morally charged scenario, $t(81)=2.94$, p<.01, but only trending towards significance in the morally neutral scenario, $t(81)=1.57$, p=.12.

Punishment

Grammatical subjects were seen as deserving more punishment than grammatical objects, $F(1,79)=19.90$, p<.001. There was also a main effect of morality such that actions in the morally charged scenario were seen as more deserving of punishment than those in the morally neutral scenario, $F(1,79)=99.86$, p<.001. Finally, there was a significant morality x syntactic position interaction $F(1,79)=10.55$, p<.01.

When the two scenarios were examined separately, there was a significant effect of syntactic position in the morally charged scenario, $t(81)=4.17$, p<.001, and also a significant effect in the morally neutral scenario, $t(81)=2.10$, p<.05.

Mediation

To further examine the relationship between syntactic position, intentionality and punishment, we conducted a meditational analysis. For this analysis, we focused specifically on judgments of the morally charged scenario and ran a series of regressions (see Figure 1). The results showed that intentionality judgments predicted punishment judgments (β =.57, p <.001). Syntactic cues predicted punishment judgments (β =-4.13,

p <.001), but when intentionality was entered into the model, the effect of syntactic

position on punishment judgment was eliminated (β =-.08, *p* =48). A Sobel test showed

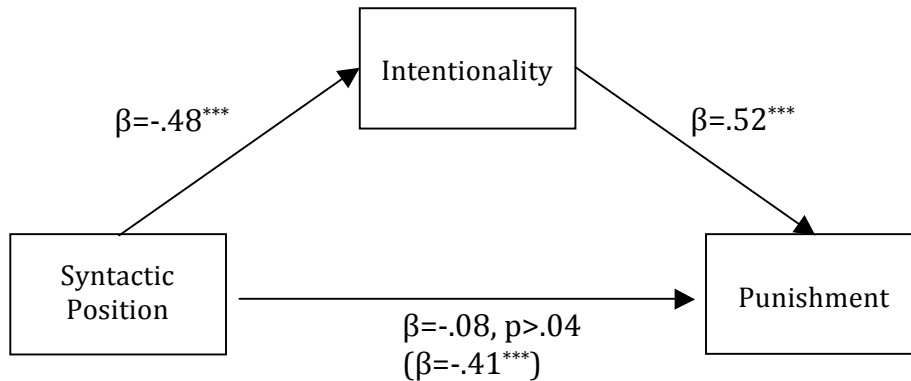that this reduction was significant, *z*=-3.75, *p* < .001.

Figure 1. Intentionality judgments mediate the impact of syntactic position on punishment judgments. *p*<.05. **p*<.01. ***p*<.001.

Discussion

This study used a manipulation that came out of the linguistics literature and then

checked for an impact on the dependent variables explored in the work of Gray and

colleagues. Across the three variables associated with agency (intentionality, responsibility,

punishment), participants consistently assigned higher ratings to the person who appeared

as the grammatical subject than to the person who appeared as the grammatical object.

For the morally charged scenario, the impact of this linguistic manipulation on the moral

judgment (punishment) was fully mediated by one of the non-moral judgments

(intentionality).

Direct objects in the morally charged scenario were viewed as being more upset

than the corresponding grammatical subject, and this difference approached statistical

significance in the morally neutral scenarios. One possibility here might be that there exists a small pre-existing difference between agents and patients in morally neutral scenarios, and that this difference is tweaked and magnified by morality.

Unsurprisingly there was more blame assigned in the morally charged compared to the morally neutral scenarios. This difference serves as a manipulation check by verifying that the participants actually understood the moral difference between what we were calling "morally charged" and "morally neutral" scenarios. More surprisingly however, the syntax manipulation influenced intuitions about punishment even in the morally neutral scenario. Though punishment ratings in the morally neutral scenario were consistently low (all means were below 2 on a 1-7 scale), there was still a statistically significant tendency whereby participants gave higher ratings to grammatical subjects than to grammatical objects.

Presumably, syntactic subjects were assigned more punishment because they were viewed as acting more intentionally, and thus deserving of blame. The mediation analysis supported this view by showing that the effect of syntactic position on punishment judgments was mediated by intentionality judgments, and that when this intermediary influence is removed, there is no significant effect of syntax on punishment.

Taken together, these results suggest that the same asymmetries in mental state attributions for agents and patients exist in both moral and neutral scenarios. This speaks in favor of the generalized view of "event dyads" which contain an intentional/responsible agent and an experiencing patient. It appears that syntax is manipulating more general aspects of theory-of-mind in all of the events. However it also influences specifically moral judgments whenever the opportunity presents itself.

Another remarkable aspect of these findings is that gender (old male vs. young

female) did not lead to any meaningful differences between groups. In other words, attributions of agentic properties like intentionality and responsibility did not differ as a function of gender/age. Similarly, attributions of patient-like properties also failed to differ as a function of gender/age. Perhaps this is because the effect of syntactic cues is so large that it simply swamped any influence of social cues.

It is worth noting that we have also replicated these findings using a different verb ("make love to") and different moral violation (that of incest) (see Supplementary Materials). This replication speaks to the generalizability and strength of the basic effects presented here.

Conclusion

This paper has been concerned with two different kinds of theories about the assignment of roles. First, there are theories about role assignments in morally charged events (as discussed by Gray and colleagues). Second, there are the theories about role assignments in events more generally (as discussed in the linguistics literature). Our aim was to understand the relationship between the roles picked out by these different kinds of theories.

The results point to a particular picture of the importance that roles like *agent* and *patient* have in people's moral judgments. It appears that people's use of these roles is not at all restricted to the moral domain. Rather, people show a quite general tendency to understand events in terms of these roles, applying them also in cases that have little moral significance. Then, when people turn their attention to morally charged events, this very general framework – which impacts judgments of intentionality and responsibility across the board – comes also to have an effect on moral judgments. In other words, it appears that the importance we see of agency and patiency within the moral domain

simply 'falls out' of the importance of these roles in people's cognition more generally.

Of course, one might complain at this point that we are simply pushing the question back a level. We began by asking why people so consistently understand morally charged events in terms of agents and patients. We then answered this first question by pointing to a broader fact about human cognition, namely, that people show a perfectly general tendency to understand events in terms of agents and patients, even when those events that have no moral significance. But now one might pose the same question again, this time at a broader level. Why is it that people show this general tendency to understand events in terms of agents and patients?

Our results do not speak directly to this deeper question, but we do want to suggest one possible hypothesis about the answer. Specifically, we propose that people's more general tendency to understand events in terms of these roles may arise from the structure of what is sometimes called 'core knowledge' (Kinzler & Spelke, 2007). That is, it may be that people have an innate preparedness to understand events in terms of these roles and that this innate preparedness then structures their thinking about events in numerous domains. People's core knowledge of role assignments could then explain the patterns we observe in their moral judgments, in their linguistic intuitions, and perhaps in many other areas besides.

References

Baker, M. 1997. Thematic roles and syntactic structure. In L. Haegeman (ed.) *Elements of Grammar.* Kluwer, Dordrecht, 73-137.

Carlson, G. & Tanenhaus M. (1988). Thematic roles and language comprehension. In W. Wilkins, (ed.), *Syntax and Semantics*, Vol. 21: Thematic Relations, Academic Press, pp. 263-89.

Dowty D.R (1991). Thematic proto-roles and argument selection. *Language* 67: 547-619.

Fillmore, C. (1968). The Case for Case. In *Universals in Linguistic Theory.* Bach. E. and Harms, R. (eds.) pp. 1-90.

Gray, K. & Wegner, D. M. (2009). Moral typecasting: Divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology*, *96*(3), 505-520.

Howard, J. A. (1984). Societal influences on attribution: Blaming some victims more than others. *Journal of Personality and Social Psychology. 47*, 494-505.

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science, 10*, 89-96.

Strickland, B., Fisher, M., Peyroux, E. & Keil, F. (2011). Syntactic biases in intentionality judgments. *Proceedings of the Thirty-Third Annual Conference of the Cognitive Science Society*.

Strickland, B., Fisher, M., Knobe, J. & Keil, F. (2012). Syntax and intentionality : An automatic link between between language and theory-of-mind. Unpublished manuscript. Yale University.