



**UvA-DARE (Digital Academic Repository)**

**How to Identify Moral Language in Presidential Speeches**

*A comparison between a social-psychological and a cognitive-linguistic approach to corpus analysis*

Renardel de Lavalette, K.Y.; Steen, G.; Burgers, C.

*Published in:*

Corpus Linguistics and Linguistic Theory

*DOI:*

[10.1515/cllt-2016-0007](https://doi.org/10.1515/cllt-2016-0007)

[Link to publication](#)

*Citation for published version (APA):*

Renardel de Lavalette, K. Y., Steen, G., & Burgers, C. (2019). How to Identify Moral Language in Presidential Speeches: A comparison between a social-psychological and a cognitive-linguistic approach to corpus analysis. *Corpus Linguistics and Linguistic Theory*, 15(2), 239-265. <https://doi.org/10.1515/cllt-2016-0007>

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

*UvA-DARE is a service provided by the library of the University of Amsterdam (<http://dare.uva.nl>)*

Kiki Y. Renardel de Lavalette\*, Gerard Steen  
and Christian Burgers

# How to identify moral language in presidential speeches: A comparison between a social-psychological and a cognitive-linguistic approach to corpus analysis

DOI 10.1515/cilt-2016-0007

**Abstract:** Lakoff (2002 [1996], *Moral politics. How liberals and conservative think*. Chicago: The University of Chicago Press) presents the Theory of Moral Politics (TMR), as based in the roles of metaphor in moral thinking in American Politics. Two distinct methods of data analysis, one social-psychological and one cognitive-linguistic, have been employed to empirically test Lakoff's assertions on moral reasoning, but have yielded different results. We applied both methods to the same corpus of speeches to determine whether they would yield similar results and could thus be considered to be equally appropriate ways of testing the presence of moral language. We show that the method affects what sort of conclusion can be drawn from research. Consequently, when testing TMR, we recommend that the corpus-linguistic method used is critically evaluated.

**Keywords:** morality, politics, corpus analysis, metaphor, political speeches

## 1 Introduction

Since the publication of *Metaphors We Live By* (Lakoff and Johnson 1980), cognitive linguists have studied metaphorical expressions in language in order

---

\*Corresponding author: Kiki Y. Renardel de Lavalette, Department of Dutch Studies, University of Amsterdam, Spuistraat 134, Amsterdam 1012 VB, Netherlands,  
E-mail: k.y.renardelavalette@uva.nl

Gerard Steen, Department of Dutch Studies, University of Amsterdam, Spuistraat 134, Amsterdam 1012 VB, Netherlands

Christian Burgers, Department of Communication Science, Vrije Universiteit Amsterdam, De Boelelaan 1081, Amsterdam 1081 HV, Netherlands

to uncover how they can reflect conventionalized ways of thinking about one domain in terms of another. Lakoff and Johnson (1980) introduced Conceptual Metaphor Theory (CMT), which posits that metaphors are not only pervasive in our everyday language (linguistic metaphors) but also in our thoughts and acts (conceptual metaphors). CMT proposes that our conceptual system is metaphorical in nature and that we understand many, typically more abstract concepts in terms of other, typically more concrete concepts.

One domain in which metaphor is thought to be of great importance is moral thinking about politics (Charteris-Black and Musolff 2003; Lakoff 2002 [1996]; Semino 2008). For this domain, Lakoff 2002 [1996] presented the Theory of Moral Reasoning (TMR), which proposes which conceptual metaphors conservatives and liberals use when thinking about political issues. TMR suggests that both conservatives and liberals think metaphorically of the state as a family, with the government fulfilling the role of a parent and the general public the role of its children. However, TMR also argues that conservatives and liberals have very different conceptions of what constitutes the ideal family. These different family ideals each come with their own subsets of metaphors for morality, and in turn, these sets of metaphors for family-based morality lead to diverging perspectives on many political issues such as abortion, welfare programs, and crime.

According to TMR, the conservative worldview centers around the Strict Father model, in which the conceptual moral metaphor Moral Strength is the core value. In this model, the father is at the head of the family and the children are raised to be self-disciplined and self-reliant by setting strict rules and implementing punishment when rules are not being followed. When abstracted to the level of politics, Strict Father morality postulates that government must enforce rule of law, and protect its citizen-children from external sources of evil. However, it must not interfere too much with the day-to-day lives of its citizen-children, since self-discipline and self-reliance are considered as requirements to be morally strong, and thus ultimately successful. This translates to policies like fixed-percentage income tax, capital punishment, and policies that support and sustain the free market economy.

The liberal worldview, by contrast, centers around the Nurturant Parent model, in which the conceptual moral metaphor Morality as Empathy is prioritized. In this model, parents and their children are equals, and children are raised to be compassionate and to have respect for others through nurturance and open communication. When abstracted to the level of politics, Nurturant Parent morality presumes the government is responsible for the well-being of its citizen-children and should be caring and involved. This translates to policies like progressive income tax, welfare programs, and environmental restrictions.

Even though proven to be popular, TMR has also received some critique. One point of criticism addressed by a number of studies is the relatively small amount of empirical support that Lakoff presents to support his theory (e. g. Bar-Lev 2007; Ohl et al. 2013). Accordingly, several studies (e. g. Cienki 2005; Deason and Gonzales 2012; Moses and Gonzales 2015; Ohl et al. 2013; Wolters 2012) have examined political discourse to test Lakoff's assertions on moral reasoning. However, at least two distinct annotation schemes have been used to answer these questions; while social-psychological scholars have identified any expression that could be considered an example of one of the two models, without considering the metaphoricity of the expression (Deason and Gonzales 2012; Moses and Gonzales 2015; Ohl et al. 2013), cognitive-linguistic scholars have analyzed texts to find metaphorical language that could be ascribed to one of the two models (Cienki 2005; Wolters 2012). These different studies yielded diverging results. Some came to the conclusion that conservatives indeed use more Strict Father language than liberals, and liberals use more Nurturant Parent language than conservatives (Cienki 2005; Moses and Gonzales 2015). Another study, however, found that both conservatives and liberals use more Strict Father language when talking about one issue, and more Nurturant Parent language when talking about some other issue (Wolters 2012). Considering these diverging results, the question arises whether these results are due to the different corpora used for the different studies, or to the fact that the different studies used different annotation schemes. The latter would imply that the various annotation schemes are not all equally appropriate ways for identifying moral language in political speeches.

In order to answer this question, this study compares one annotation scheme based on the social-psychological approach (Moses and Gonzales 2015) to one annotation scheme with a cognitive-linguistic approach (Wolters 2012) by applying them to the same corpus of twenty speeches, ten of which were delivered by George W. Bush and ten of which were delivered by Barack Obama. The results of the two annotation schemes are compared to determine whether they yield similar results and could thus be considered to be equally appropriate ways of testing the presence of moral language according to TMR.

## 1.1 Empirical studies with a social-psychological approach

First, various studies used a social-psychological approach to identify language that could be ascribed to either the Strict Father or the Nurturant Parent model

(e.g. Deason and Gonzales 2012; Moses and Gonzales 2015; Ohl et al. 2013). In these social-psychological studies, any expression that could be considered an example of the Strict Father or Nurturant Parent models was identified, without considering its metaphoricity. However, CMT proposes that conceptual metaphors are reflected in language through linguistic metaphors. The social-psychological approach thus raises the question whether metaphors in thought can be reflected by non-metaphorical expressions in language. According to cognitive linguist Cienki (2005), non-metaphorical expressions should indeed reflect the broader, nonpropositional logic of the proposed models for morality, and could therefore either offer support or contradict the idea that conservatives and liberals think in terms of these models.

These social-psychological studies first construed a list of Strict Father and Nurturant Parent categories based on Lakoff's (2002 [1996]) description of the two models, but they did so in slightly different ways. Deason and Gonzales (2012) generated a classification system based on the descriptions of the two family models and the moral metaphors by Lakoff (2002 [1996]), and modified this scheme based on free responses which were obtained in a previous study in which participants described the characteristics of the ideal family and the ideal government (Deason et al. 2008). This resulted in eleven categories for the Strict Father model, such as Work Ethic, Self-Interest, Right and Wrong, and Moral control, and fifteen categories for the Nurturant Parent model, of which some examples are Nurturance, Helpfulness, Connectedness, and Open-mindedness.

Ohl et al. (2013) based their taxonomy solely on the descriptions of the Strict Father and Nurturant parent models, which resulted in eight subordinate categories linked to Strict Father reasoning, such as Danger comes from human agents, Opposes illegitimate authority, and emphasizes moral strength, and eight subordinate categories linked to Nurturant Parent reasoning, such as Danger comes from bad environment, Encourages leading by example, and Encourages self-nurturance.

Moses and Gonzales (2015) based their taxonomy on the descriptions of the two moral-metaphorical frameworks that follow from the conceptions of the two family models. Their taxonomy comprises a number of subordinate metaphors such as Moral Strength and Morality as Nurturance, and a number of subcategories subsumed under each general orientation, such as discipline, competition, cooperation and taking care of others, which resulted in eight Strict Father categories and six Nurturant Parent categories. These studies thus did not use the categories as described by Lakoff (2002 [1996]).

Furthermore, when we compare the categories of the three studies to each other, we see that their taxonomies do not correspond one-to-one. All three used

different labels for their categories, and also gave them different descriptions. Furthermore, the studies used different numbers of categories and subcategories, so what is captured within one category in one study is captured within three different categories in another study. However, all three studies have at least one category concerning Morality as Strength, and one concerning Morality as Nurturance.

The analytical procedures as applied by Deason and Gonzales (2012) and Moses and Gonzales (2015) are more or less similar. In both studies, the corpora were divided into either “phrases, clauses, or whole sentences as units of analysis, all of which were held to express a single complete thought” (Deason and Gonzales 2012: 258; Moses and Gonzales 2015: 383). Each unit of analysis was then coded for the presence of linguistic instantiations captured by the Strict Father and Nurturant Parent taxonomies. The data analyzed by Deason and Gonzales (2012) consisted of written transcripts of the speeches at National Conventions of Democrats Barack Obama and Joe Biden, and Republicans John McCain and Sarah Palin. Moses and Gonzales (2015) studied written transcripts of 202 presidential campaign advertisements. Ohl et al. (2013) applied a different coding method. In this study, complete presidential campaign advertisements were coded for the presence of each of the sixteen categories that they distinguish. Nevertheless, these studies are all similar in the fact that they do not consider the metaphoricality of expressions in texts.

The number of expressions representing the Strict Father or Nurturant Parent models found in these studies was relatively high. The results, however, differed between the studies. Deason and Gonzales (2012) found that Democrats used significantly more expressions belonging to Nurturant Parent themes than Republicans, but Republicans did not significantly use more Strict Father themes than Democrats. Furthermore, while Democrats used twice as many Nurturant Parent expressions than Strict Father utterances, Republicans used similar numbers of expressions belonging to the Strict Father model as to the Nurturant Parent model. The results of Ohl et al.’s (2013) and Moses and Gonzales’ (2015) studies show almost the opposite. Ohl et al. (2013) found that that Republicans outpace Democrats in their overall use of moral reasoning. Furthermore, while Republicans made significantly more use of Strict Father expressions than Democrats, Democrats did not significantly make more use of Nurturant Parent expressions than Republicans. Moses and Gonzales’ (2015) study showed that Republicans used more language belonging to the Strict Father model, and that Democrats used more Nurturant Parent language. However, Republicans made almost no use of the Nurturant Parent paradigm, while Democrats used a fair amount of Strict Father expressions.

## 1.2 Empirical studies with a cognitive-linguistic approach

Second, a number of cognitive-linguistic scholars applied annotation schemes that differed from those of the social-psychological scholars (e. g. Cienki 2005; Wolters 2012). In these studies, texts were examined for the presence of metaphorical language reflecting the conceptual metaphors for morality as described by Lakoff (2002 [1996]).

Cienki (2005) made a compilation of the metaphors listed in *Moral Politics*, in which Lakoff (2002 [1996]) outlines the Strict Father and Nurturant Parent models. Subsequently, Cienki (2005) analyzed written transcripts of three 90-minute televised debates between George W. Bush and Al Gore, from the October 2000 Presidential elections. Cienki (2005) coded for the presence of words and phrases that were deemed to express any of the metaphors present in the compilation. In particular, he focused on terms used in the debates that could be related to any of the source domains present in the list of metaphors for morality, and which were deemed to be used to talk about the relevant target domains, and on this basis expressions were classified as either Strict Father or Nurturant Parent expressions. However, since this study was conducted in 2005, a systematic and reliable method for the identification for linguistic metaphors such as the Metaphor Identification Procedure (MIP) (Pragglejaz Group 2007) or the Metaphor Identification Procedure – Vrije Universiteit (MIPVU) (Steen et al. 2010) did not yet exist, nor did Cienki (2005) explicitly state how he decided whether an expression was considered to be metaphorical or not. Furthermore, Cienki (2005) did not use one consistent unit of analysis and marked expressions ranging in length from individual words to syntactic phrases on or below sentence level.

Wolters (2012) applied a different method for the identification of metaphorical expressions reflecting the conceptual metaphors for morality as described by Lakoff (2002 [1996]). He analyzed a corpus comprising ten speeches given by George W. Bush and ten by Barack Obama, half of which were on Health Care and the other half on the war in Iraq by selecting a number of representative nouns, verbs, and adjectives for each of the eight prioritized metaphors for morality per family model as described by Lakoff (2002 [1996]). Thus, for the metaphor Moral Strength, Wolters (2012) compiled a list of nouns, verbs, and adjectives related to the source domain strength, such as *stability*, *to strengthen*, and *courageous*. For the metaphor Morality as Nurturance, Wolters (2012) compiled a similar list with lexical items reflecting the source domain nurturance, including *harbor*, *to attend to*, and *to serve*. In this way, Wolters (2012) follows Koller (2008), who used this technique for studying two corpora consisting of articles on marketing and sales, and

mergers and acquisition, by selecting a number of lexical items representing dominant conceptual metaphors in economic discourse. Next, Wolters (2012) used the MIPVU procedure (Steen et al. 2010) to decide whether the instances of the listed lexical items in his corpus were used metaphorically or not. Lastly, Wolters (2012) decided for the metaphorically used lexical items whether they pertained to morality or not. However, Wolters (2012) only gives some examples of expressions thought to refer to morality and some that did not, and does not explicate how he defines morality or how he decided whether an expression refers to morality.

Both Cienki (2005) and Wolters (2012) found only a small number of metaphorical expressions that could be ascribed to either the Strict Father or the Nurturant parent model, and their results diverge. Cienki's (2005) results showed that George W. Bush did indeed use more metaphorical expressions belonging to the Strict Father model than to the Nurturant Parent model, and Al Gore used more Nurturant Parent metaphorical expressions, just as predicted. Wolters (2012), by contrast, found that within the topic of Health Care both George W. Bush and Barack Obama used more expressions belonging to the Strict Father model than the Nurturant Parent model, while within the topic of the war in Iraq both presidents used more expressions belonging to the Nurturant Parent than to the Strict Father model. Hence, Wolters (2012) concluded that there is no correlation between president and metaphoricity based on the Strict Father of Nurturant Parent model, but that there is instead a correlation between topic of speech and morality model.

### 1.3 Current study

Given the different annotation schemes and the diverging outcomes of the different studies, it remains to be seen whether the two different ways of data coding actually both measure the same concepts from TMR. Therefore, the current study compares one annotation scheme from each approach to each other, by applying them to the same corpus of speeches, to test whether they yield comparable results. When generating congruent results, both annotation schemes are supposedly equally suitable ways for measuring the presence of moral language belonging to either the Strict Father or Nurturant Parent model in authentic political discourse. Thus, this study answers the question whether the social-psychological and cognitive-linguistic annotation schemes yield similar or divergent results when applied to the same corpus of political speeches, and can consequently be considered to be equally appropriate ways of testing the presence of moral language in political discourse.



## 2 Methods

### 2.1 Data

The data used for this study has been selected and analyzed by Wolters (2012). He chose these speeches because of the relevance of their topics to US politics, and because he assumed that these two topics would invoke a high level of morality, and thus moral language due to their highly value-laden character. The data was used again for this study, since the goal of this study is to compare two different annotation schemes aimed at identifying moral language in political discourse. Besides, Wolters (2012) did not explain how he decided whether a metaphorical expression referred to morality, which is an aspect that we improved upon in this study.

The data comprises twenty speeches. Ten of these speeches were given by George W. Bush, of which five were on Health Care and five on the war in Iraq. The other ten speeches were delivered by Barack Obama, and again half of these speeches were on Health Care and the other half on the war in Iraq. All speeches were delivered between the years 2001 and 2010. It should be noted that not all of the ten speeches delivered by Obama are from his time as President of the United States. Some of them were delivered when Obama was still a Senator for the state of Illinois.

The total corpus comprises approximately 60,000 words. This number of words is more or less evenly divided over the four sets of speeches. Thus, the five speeches on Health Care by Obama and by Bush each comprise approximately 15,000 words, as do the five speeches on the war in Iraq by either president. Table 1 gives an overview of the exact number of words per sub-corpus.

**Table 1:** The number of words per sub-corpus.

	George W. Bush	Barack Obama	Total
Iraq War corpus	15,493	15,098	30,591
Health Care corpus	13,671	16,652	30,323
Total number of words	29,164	31,750	60,914

As Wolters (2012) notes, speeches are often not written by the person delivering the speech, and do thus not necessarily represent the thoughts and beliefs of only the speaker with regard to the issue in question. However, even though the speeches might have been written by others than Bush or Obama

themselves, they do represent the views of the office of the American president and should therefore contain manifestations of the two moral political models, in line with TMR (Cienki 2005).

## 2.2 Method 1: A social-psychological approach to data coding

The social-psychological annotation scheme applied to the corpus is based on the one as used by Moses and Gonzales (2015). This annotation scheme is considered most appropriate for the purpose of this study because of two main reasons. First, of the three studies, their taxonomy stays closest to the descriptions of the two models as given by Lakoff (2002 [1996]), especially compared to the study of Deason and Gonzales (2012) who supplemented their coding taxonomy with input from participants in a previous study. Second, Ohl et al. (2013) used complete advertisements as units of analysis, and their coding taxonomy is adjusted accordingly. Therefore, the annotation scheme of Moses and Gonzales (2015) is also better suited for comparison than that of Ohl et al. (2013).

In order to improve reliability, some adjustments were made to the original annotation scheme as presented by Moses and Gonzales (2015). First, Moses and Gonzales (2015: 383) defined the unit of analysis as “phrases, clauses, or whole sentences that expressed a single complete thought (text elements)”. This method for dividing a text into units of analysis proved to be problematic due to the lack of a straightforward operationalization. Therefore, the unit of analysis is here defined as a discourse unit, where restrictive clauses belong to the main sentence as a dependent clause and together they form one unit of analysis, while expanding clauses are treated as independent units of analysis (following Steen 2005). Two coders applied this method for dividing a corpus into units of analysis to a subsection of the data in order to assess whether they would do so in a similar way. In 199 instances the two coders agreed on the division of the texts into separate units of analysis. In eight instances the first coder divided the text into separate units of analysis while the second coder did not, and in ten instances the second coder separated the texts into different units of analysis where the first coder did not. The margin of error is less than 10 % and thus this method of dividing a text into units of analysis is considered reliable.

As for coding the units of analysis into moral political categories, Moses and Gonzales (2015) based their taxonomy for the Strict Father and Nurturant Parent categories on Lakoff's (2002 [1996]) descriptions of the two moral frameworks belonging to the two different family models. The Strict Father categories that

Moses and Gonzales (2015) formulated include Morality as Strength, Morality as Self-Discipline, Competition, Moral Authority, Moral Contagion, Tough Love, Self-Reliance, and Other SF ideas. The Nurturant Parent categories include Morality as Nurturance, Responsibility for others, Cooperation, Openness, Involved and Responsible Authority, and Other NP ideas. Furthermore, they added an “uncodeable” category to capture ideas irrelevant to the Strict Father and Nurturant Parents models.

However, from the descriptions of the various categories as given by Moses and Gonzales (2015), it is not evident how they determined whether a unit of analysis was an expression of Moral Politics and how one of the Strict Father or Nurturant Parent categories could subsequently be assigned to the relevant units of analysis. They gave a handful of examples and these suggest that they did not only annotate utterances referring to Moral Politics, but any expression that could somehow be explained by TMR. For instance, they use the following utterance as an example of the Nurturant Parent category Openness: *With new ideas to make life better*. However, in this example it is not stated that new ideas are moral imperatives in order for life to get better, it only implies that new ideas will be used to make life better. To solve this issue, some adjustments were made, and the following steps were added.

First, for each unit of analysis, we determined whether it addresses a political topic or not. An utterance is considered to be political when addressing an issue that is on the public agenda, and that involves a considerable number of individuals within a society. An example of such an expression is the following sentence uttered by Barack Obama in a speech on Health Care: *Now, the proposal I put forward gives Americans more control over their health insurance and their health care by holding insurance companies more accountable*. In this utterance, Obama talks about a topic that is on the public agenda (i. e. health insurance), and that concerns the American people. The next sentence, coming from a speech given by Bush on the war in Iraq, is an example of an utterance that does not refer to a political topic: *Thank you for that very gracious and warm Cincinnati welcome*. In this example, Bush thanks his audience, and he is thus not addressing a topic that is on the public agenda, or that concerns a considerable number of people in society. To assess reliability, a second coder coded a sub-corpus comprising 300 units of analysis. This resulted in a Cohen's Kappa of 0.63 (“substantial agreement”; Landis and Koch 1977).

Second, for the units of analysis that were deemed to refer to politics, we decided whether these particular units of analysis were an expression of morality or not. An expression is considered moral when referring to conceptions of right or wrong behavior, and good or bad actions or characteristics of people or governments. These utterances can be either descriptive or

evaluative. An example of an expression referring to moral politics is the following sentence from a speech on Health Care, given by Bush: *Just as we addressed discrimination based on race, we must now prevent discrimination based on genetic information.* In this utterance, Bush addresses a form of behavior that he believes to be wrong. An utterance referring to politics that is not considered to be moral is the following example from a speech on the war in Iraq, given by Obama: *To understand where we need to go in Iraq, it is important for the American people to understand where we now stand.* In this example, Obama talks about what sort of action is needed to achieve a certain goal, not about what kind of behavior is right or wrong. To determine inter-coder-reliability for this step, the second coder coded a subcorpus of 200 units of analysis. Cohen's Kappa was 0.67 ("substantial agreement"; Landis and Koch 1977).

If a unit of analysis refers to a political topic and is also coded as a moral expression, then the unit of analysis is considered an expression of Moral Politics, and can be further analyzed according to the taxonomy of Moses and Gonzales (2015). Thus, subsequently, each of the expressions that refers to a political topic and is considered to be a moral expression was assigned one, none, or several of the Strict Father or Nurturant Parent categories as formulated by Moses and Gonzales (2015).<sup>1</sup> For further calculations and interpretations, it was only taken into account whether an expression could be subsumed under the Strict Father and/ or the Nurturant Parent models, irrespective of the precise categories that were assigned to an utterance. Inter-coder-reliability was calculated for expressions that were considered to be subsumed under the Strict Father model and for expressions that were considered to be subsumed under the Nurturant Parent model separately. The Cohen's Kappa for the Strict Father coding was 0.86 ("almost perfect agreement", Landis and Koch 1977), and for the Nurturant Parent coding it was 0.70 ("substantial agreement", Landis and Koch 1977). The two sub-corpora coded by the second coder to determine the inter-coder-reliability both comprised 200 units of analysis.

Another adjustment that was made is that the Moses and Gonzales' (2015) categories Other SF ideas and Other NP ideas were omitted, since no explicit description of what these two categories constituted was given. Lastly, the term "uncodeable" was changed into "unrelated" to avoid confusion as to whether the annotation scheme is indecisive with regard to the units of analysis in question or whether the unit of analysis in question is not considered to be an expression of Moral Politics.

---

<sup>1</sup> Examples from our corpus of the Strict Father and Nurturant Parent categories as formulated by Moses and Gonzales (2015) can be found in Tables 2 and 3.

**Table 2:** Strict Father categories in the social-psychological annotation scheme.

Strict Father Categories	Description	Examples
SF1: Morality as Strength	Self-control; toughness; being strong against immorality, evil or adversity.	“Yet our coalition is strong, our efforts are focused and unrelenting, and no power of the enemy will stop Iraq’s progress.” (Bush)
SF2: Morality as Self-Discipline	Discipline; determination; motivation.	“(.) and despite setbacks and battlefield defeats, we did not waver in freedom’s cause.” (Bush)
SF3: Competition	Competition is moral and ensures success. Also the importance of trying to compete successfully.	“Competition has always been a catalyst for American innovation, and now should be no different.” (Obama)
SF4: Moral Authority	Leaders must, should, or do have the authority to lead.	“When America leads with principle and pragmatism, hope can triumph over fear.” (Obama)“But that is not an excuse for those of us who were sent here to lead.” (Obama)
SF5: Moral Contagion	Right and wrong used in an absolute sense; evil can infect moral others.	“ Without American leadership, these threats will fester.” (Obama)
SF6: Tough love	Hardship and failure as “tough love” are good for people; too much help harms people.	–
SF7: Self-Reliance	Advocating personal responsibility and independence rather than government intervention.	“America has wisely resisted efforts to place medicine and health care into the hands of central planners.” (Bush)

**2.3 Method 2: A cognitive-linguistic approach to data coding**

The second annotation scheme that was adopted for the analysis of the corpus was the one as applied by Wolters (2012), which was based on a technique of corpus analysis used by Koller (2008) and on the MIPVU procedure (Steen et al. 2010) for the analysis of linguistic metaphors. This annotation scheme was chosen over the annotation scheme of Cienki because Wolters’

**Table 3:** Nurturant Parent categories in the social-psychological annotation scheme.

Nurturant Parent Categories	Description	Examples
NP1: Morality as Nurturance	Nurturing others; love, and kindness.	“That large-heartedness – that concern and regard for the plight of others – is not a partisan feeling.” (Obama)
NP2: Responsibility for others	Helping and providing direct care for others as moral imperatives, especially those less fortunate and vulnerable.	“And you have borne an enormous burden for your fellow citizens, while extending a precious opportunity to the people of Iraq.” (Obama)
NP3: Cooperation	Working together is moral and ensures success.	“And we need to support the forces of moderation in the Islamic world, so that alliances of convenience mature into friendships of conviction.” (Obama)
NP4: Openness	Taking the others perspective; understanding, and openness to new and different ideas.	“Government must be willing to change with the times.” (Bush)
NP5: Involved, Responsible Authority	Authority figures have the responsibility to be involved and instrumental on behalf of those with less power or authority.	When Congress finally rose above politics and fulfilled its duty to America’s seniors, it showed what’s possible in Washington, D.C. (Bush)

method of metaphor identification, MIPVU (Steen et al. 2010), is clear-cut and its reliability has been substantially tested, while Cienki’s method for metaphor identification is not clearly described and thus not as easily replicable as Wolters’ method. Just like the social-psychological annotation scheme, this annotation scheme was also slightly adjusted to make it more reliable than the original annotation scheme as applied by Wolters (2012).

For both the Strict Father and the Nurturant Parent model, Wolters (2012) identified the eight prioritized conceptual metaphors for morality as described by Lakoff (2002 [1996]). For the Strict Father model these prioritized metaphors for morality are: Moral Strength, Moral Authority, Moral Essence, Moral Wholeness, Moral Order, Moral Boundaries, Moral Purity, and Moral Health. The eight prioritized metaphors for morality belonging to the Nurturant Parent model are: Morality as Empathy, Morality as Nurturance, Moral Self-nurturance, Morality as Social Nurturance, Morality as Happiness, Morality as Self-development, Morality as Fair Distribution, and Moral Growth.

Following the example of Koller (2008), Wolters then identified a number of representative lemmas and lexemes, henceforward referred to as lexical items, for each of the source domains of these prioritized conceptual metaphors for morality of the two models. The lexical items that Wolters (2012) chose were all taken from Lakoff's book *Moral Politics* (2002 [1996]), and from a study concerning TMR by Cienki (2005). The total number of lexical item types that represented the Strict Father Moral categories was 91, of which 40 were nouns, 26 verbs, and 25 adjectives. The total number of lexical item types representing the Nurturant Parent Moral categories was 94, of which 40 were nouns, 40 verbs, and 14 adjectives.

Next, The MIPVU procedure (Steen et al. 2010) was applied to all tokens of the listed lexical items in the corpus to decide whether the lexical units in question were used metaphorically or not (Cohen's Kappa = 0.68, "substantial agreement"; Landis and Koch 1977). A detailed description of how MIPVU is to be applied can be found in *A Method for Linguistic Metaphor Identification: From MIP to MIPVU* (Steen et al. 2010).

The last step in Wolters' (2012) annotation scheme concerns the decision whether the target domain of the metaphorically used lexical units is morality or not. However, Wolters (2012) does not explain how he made this decision. He gives a number of examples from the corpus that he considers instances of metaphorically used lexical items with morality as target domain, and a number that do not have morality as a target domain, but he does not provide any instructions on which he based these decisions. In order to solve this issue, the same solution as presented for method 1 was applied to this annotation scheme. This means that the metaphorically used lexical items that occur in the units of analysis that was annotated as being expressions of Moral Politics in method 1 are considered as expressions in which morality is the target domain of the metaphorical expression, and are thus representative for the moral priority in question. The Cohen's Kappa's for the two steps taken to decide whether an expression was referring to Moral Politics was 0.63 ("substantial agreement", Landis and Koch 1977) for determining whether an utterance was about politics, and 0.67 for determining whether an expression referred to morality ("substantial agreement", Landis and Koch 1977).<sup>2</sup>

Because of the adjustments made to the cognitive-linguistic annotation scheme, the results yielded by this study differ from those of Wolters (2012).

---

<sup>2</sup> Examples from our corpus of the Strict Father and Nurturant Parent categories as formulated by Wolters (2012) can be found in Tables 4 and 5.

Table 4: Strict Father categories in the cognitive-linguistic annotation scheme.

Strict Father Categories	Lexical items	Examples
Moral Strength	Stability, Stable, Iron, Strength, to Strengthen, Vigor, Vigorous, Courage, Courageous, Power, Powerful	“Helping construct a <u>stable</u> democracy after decades of dictatorship is a massive undertaking.”(Bush)
Moral Order	Balance, to Balance, Rank, to Rank, Status, Hierarchy, Hierarchical, Reign, to Reign, to Dominate, Dominant	“And John McCain wants to give insurance companies free <u>reign</u> to avoid them.” (Obama)
Moral Authority	Grip, to Grip, Command, to Command, Force, to Force, Leader, to Lead, Leading, Demand, to Demand, Rule, to Rule	“This will require American <u>force</u> levels.”(Bush)
Moral Boundaries	Edge, Separation, to Separate, Limit, to Limit, Transgression, to Transgress, to Close, Confinement, to Confine, Confined	“We will place a <u>limit</u> on how much you can be charged for out-of-pocket expenses, because in the United States of America, no one should go broke because they get sick.”(Obama)
Moral Essence	Core, Value, to Value, Vision, Confidence, Confident, Honesty, Honest, Patience, Patient	“The <u>core</u> leadership of Al Qaeda has a safe-haven in Pakistan.”(Bush)“I have a different <u>vision</u> .”(Obama)
Moral Purity	Dirt, Dirty, Spotless, Transparency, Transparent, to Clean, Clean, Genuine, Purity, to Purify, Pure	“There ought to be <u>transparency</u> in medicine.” (Bush)
Moral Wholeness	Damage, to Damage, Fragility, Fragile, Join, to Join, Wholeness, Whole, Unity, to Unite, United, Complete	(..) and they must step up their support for Iraq’s <u>unity</u> government.”(Bush)
Moral Health	To Freshen, Fresh, Cure, to Cure, to Heal, Recovery, to Recover, Condition, Fitness, Fit	“(..) and help it [Iraq] <u>recover</u> from decades of brutal tyranny.”(Obama)

According to Wolters (2012), Bush used 108 lexical items that can be considered to be reflections of the moral metaphors described by Lakoff (either belonging to the Strict Father or Nurturant Parent model), and Obama used 170. This study, however, found that Bush used 198 lexical items as metaphors for morality, and Obama used 201.



**Table 5:** Nurturant Parent categories in the cognitive-linguistic annotation scheme.

Nurturant Parent Categories	Lexical items	Examples
Morality as Empathy	Interest, Sense, to Sense, Understanding, to Understand, Feeling, to Feel, Benevolence, Benevolent, Grace, Gracious	“It’s a fundamental improvement and it makes a lot of <u>sense</u> .”(Bush)
Morality as Nurturance	To Feed, Harbor, to Harbor, to Underpin, To attend to, Aid, to Aid, to Serve	“Thanks for <u>serving</u> our country.”(Bush)
Moral Self-Nurturance	Gain, to Gain, Sacrifice, to Sacrifice, Reward, to Reward, Aspiration, to Aspire, Determination, to Determine, Determined, Practice, to Practice	“We respect your sovereignty and the tremendous <u>sacrifices</u> you have made for your country.”(Obama)
Morality as Social Nurturance	Maintenance, to Maintain, Satisfaction, to Satisfy, Satisfactory, Tie, to Tie, Solidarity, Support, to Support, to Tolerate, Tolerant	“Seniors who receive drug coverage from a former employer or union can count on new <u>support</u> from Medicare to help them keep their good benefits.”(Bush)
Morality as Happiness	Energy, Energetic, Glow, to Glow, Bright, to Brighten, Joy, Joyful, Cheer, to Cheer, Cheerful, Glad	-
Morality as Self-Development	Design, to Design, Maturity, to Mature, Mature, Enrichment, to Enrich, Structure, to Structure, Perfection, to Perfect, Perfect, Development, to Develop, Developmental	“And we need to support the forces of moderation in the Islamic world, so that alliances of convenience <u>mature</u> into friendships of conviction.”(Obama)
Morality as Fair Distribution	Expansion, to Expand, Spread, to Spread, Split, to Split, Equality, to Equate, Division, to Divide, Share, to Share	“(.) and we’re going to make sure that the wealthiest Americans pay their fair <u>share</u> .”(Obama)
Moral Growth	Root, to Root, Flourish, to Flourish, Rise, to Rise, Growth, to Grow, to Originate, to Prosper	“When congress finally <u>rose</u> above politics and fulfilled its duty to America’s seniors, it showed what’s possible in Washington D.C.”(Bush)

### 3 Results

#### 3.1 Results: Method 1

Following Moses and Gonzales’ (2015) annotation scheme, the twenty speeches delivered by Bush and Obama contained a considerable number of discourse units that could be ascribed to one of the two moral political models described by Lakoff (2002 [1996]). 38.77 % of all discourse units in the Bush corpus are captured by either the Strict Father or Nurturant Parent categories, compared to 45.67 % of all discourse units in the Obama corpus (see Table 6).

**Table 6:** The number of discourse units that reflect the SF or NP models, or that are unrelated, in the Bush and the Obama corpus, by applying the social-psychological annotation scheme.

	Bush				Obama			
	SF expressions	NP expressions	Unrelated items	Total	SF expressions	NP expressions	Unrelated items	Total
Iraq War	389	207	602	1,198	391	234	435	1,060
Health Care	114	207	846	1,167	146	283	819	1,248
Total	503	414	1,448	2,365	537	517	1,254	2,308

In the speeches delivered by Bush, 21.27 % (503 out of 2,365) of all discourse units are part of the Strict Father categories, and 17.51 % (414 out of 2,365) of the Nurturant Parent categories. In Obama’s speeches, 23.27 % (537 out of 2,308) of all discourse units are captured by the Strict Father categories, and 22.40 % (517 out of 2,308) by the Nurturant Parent categories.

In order to test whether one president used more Strict Father and Nurturant Parent language than the other, we compared the two presidents in a series of chi-square tests. Besides, because including effect sizes is a necessary practice to determine the impact of the results found (e. g. Kotrlik and Williams 2003), we used the Odds Ratio as effect size metric.

Overall, we found no significant relation between President and the use of Strict Father expressions (odds ratio = 1.12, 95 %CI = [0.98, 1.29],  $\chi^2$  (1) = 2.70,  $p$  = 0.10), but we did find a significant relation between President and the use of Nurturant Parent expressions: based on the odds ratio, the odds of Obama using a Nurturant Parent expression was 1.36 times (95 %CI = [1.18, 1.57]) larger than Bush ( $\chi^2$  (1) = 17.54,  $p$  < 0.05).

Within the speeches on the war in Iraq, we found a significant relation between President and the use of Strict Father expressions: based on the odds ratio, the odds of Obama using a Strict Father expression was 1.22 times (95 % CI = [1.02, 1.45]) larger than Bush ( $\chi^2(1) = 4.85, p < 0.05$ ). We also found a significant relation between President and the use of Nurturant Parent expressions: based on the odds ratio, the odds of Obama using a Nurturant Parent expression was 1.36 times (95 % CI = [1.10, 1.67]) larger than Bush ( $\chi^2(1) = 8.23, p < 0.05$ ). As for the topic of Health Care, we found no significant relation between President and the use of Strict Father expressions (odds ratio = 1.22, 95 % CI = [0.94, 1.59],  $\chi^2(1) = 2.34, p = 0.13$ ), but we did find a significant relation between President and the use of Nurturant Parent expressions: based on the odds ratio, the odds of Obama using a Nurturant Parent expression was 1.36 times (95 % CI = [1.11, 1.66]) larger than Bush ( $\chi^2(1) = 9.09, p < 0.05$ ).

### 3.2 Results: Method 2

Table 7 shows an overview of the numbers of metaphorical lexical items reflecting the Strict Father and Nurturant Parent models in the speeches by Bush and Obama on the Iraq War and on Health Care.

**Table 7:** The number of lexical items that reflect the SF or NP models, or that are unrelated, in the Bush and the Obama corpus, by applying the cognitive-linguistic annotation scheme.

	Bush				Obama			
	SF expressions	NP expressions	Unrelated items	Total	SF expressions	NP expressions	Unrelated items	Total
Iraq War	79	58	15,356	15,493	85	78	14,934	15,098
Health Care	21	40	13,610	13,671	21	17	16,614	16,652
Total	100	98	28,966	29,164	106	95	31,549	31,750

In the speeches delivered by Bush, the total number of lexical items that potentially reflect the Strict Father or Nurturant Parent metaphorical moral systems that were found by applying the cognitive-linguistic annotation scheme as constructed by Wolters (2012) is 509. This means that 1.75 % of the total number of lexical items present in the speeches of Bush are lexical items that Wolters (2012) listed as potentially reflecting the moral metaphors of the Strict Father or the Nurturant Parent models as described by Lakoff (2002 [1996]).

Of these 509 lexical items, 265 belong to the Strict Father model, and 245 belong to the Nurturant Parent model. However, only 100 of the 265 (37.74 %) lexical items were deemed to actually reflect metaphors for morality belonging to the Strict Father model, and only 98 of the 245 (40 %) lexical items reflected the metaphors for morality belonging to the Nurturant parent model.

In the speeches given by Obama, the total number of lexical items potentially reflecting the Strict Father or Nurturant Parent metaphorical moral systems is 438, which is 1.38 % of the total number of lexical items making up the Obama corpus. 220 of the 438 lexical items belong to the Strict Father model, and 218 to the Nurturant Parent model. However, only 106 of the 220 (48.18 %) lexical items were considered to be an actual reflection of the metaphors for morality belonging to the Strict Father model, and 95 of the 218 (43.58 %) lexical items reflect the metaphors for morality belonging to the Nurturant parent model.

Just as for the first annotation scheme, a series of chi-square tests were conducted, and odds ratios were calculated, to test whether Bush used a similar or different amount of moral language than Obama. Overall, we found no significant relation between President and the use of Strict Father expressions (odds ratio = 1.03, 95 %CI [0.78, 1.35],  $\chi^2(1) = 0.04$ ,  $p = 0.85$ ), nor a significant relation between President and the use of Nurturant Parent expressions (odds ratio = 1.12, 95 %CI [0.85, 1.49],  $\chi^2(1) = 0.65$ ,  $p = 0.42$ ).

For the speeches on the topic of the war in Iraq, we also found no significant relation between President and the use of Strict Father expressions (odds ratio = 1.11, 95 %CI [0.81, 1.50],  $\chi^2(1) = 0.40$ ,  $p = 0.53$ ), and no significant relation between President and the use of Nurturant Parent expressions (odds ratio = 1.38, 95 %CI [0.98, 1.94],  $\chi^2(1) = 3.50$ ,  $p = 0.06$ ). Within the speeches on Health Care, we found no significant relation between President and the use of Strict Father expressions (odds ratio = 1.22, 95 %CI [0.67, 2.23],  $\chi^2(1) = 0.41$ ,  $p = 0.52$ ), but we did find a significant relation between President and the use of Nurturant Parent expressions: based on the odds ratio, the odds of Bush using a Nurturant Parent expression was 2.87 times (95 %CI = [1.63, 5.07]), larger than Obama ( $\chi^2(1) = 14.52$ ,  $p < 0.05$ ).

### 3.3 Results: Method 1 and 2 compared

Although the two distinct methods under study are both aimed at empirically testing TMR by analyzing political discourse, and would thus ideally lead to similar results, they yielded diverging results, when applied to the same corpus.

First, when analyzing the complete set of speeches, the results of the social-psychological approach to data analysis shows that Obama and Bush used a similar number of expressions belonging to the Strict Father model, but that Obama used more Nurturant Parent language than Bush (odds ratio = 1.36, 95 % CI = [1.18, 1.57],  $\chi^2 (1) = 17.54$ ,  $p < 0.05$ ).

The results of the cognitive-linguistic approach, however, indicate that Bush and Obama used a similar number of lexical items ascribed to the Strict Father model as well as a similar number of lexical items belonging to the Nurturant Parent model.

The two methods also yield diverging results when considering the two topics of the speeches separately. As for the speeches on the war in Iraq, the results of the social-psychological method show that Obama used more expressions belonging to the Strict Father model (odds ratio = 1.22, 95 % CI = [1.02, 1.45],  $\chi^2 (1) = 4.85$ ,  $p < 0.05$ ) as well as to the Nurturant Parent model (odds ratio = 1.36, 95 % CI = [1.10, 1.67],  $\chi^2 (1) = 8.23$ ,  $p < 0.05$ ) than Bush. Divergently, the cognitive-linguistic method indicates that Bush and Obama used a similar number of expressions subsumed under the Strict Father model, as well as under the Nurturant Parent model.

Lastly, when considering the speeches on Health Care, the results of the two methods both indicate that Bush and Obama use a similar number of expressions subsumed under the Strict Father model. However, they yield diverging results regarding the use of Nurturant Parent language. The social-psychological method finds that Obama used more Nurturant Parent language than Bush (odds ratio = 1.36, 95 % CI = [1.11, 1.66],  $\chi^2 (1) = 9.09$ ,  $p < 0.05$ ), while the cognitive-linguistic method contrastingly finds that Bush used more expressions ascribed to the Nurturant Parent model than Obama (odds ratio = 2.87, 95 % CI = [1.63, 5.07],  $\chi^2 (1) = 14.52$ ,  $p < 0.05$ ).

To conclude, although we find some similarities in the results yielded by applying the two distinct methods of data analysis to the same corpus, we also find a considerable number of differences. Table 8 gives a summary of the statistical results obtained by applying the social-psychological method and the cognitive-linguistic method to the corpus of political speeches. This implies that although the two methods of data analysis are both aimed at empirically testing TMR by analyzing political discourse, they do not equivalently measure this phenomenon.

## 4 Discussion and conclusion

This study set out to compare two divergent annotation schemes for testing TMR in order to find out whether they would yield similar results when being

**Table 8:** Summary of the statistical results obtained by applying the social-psychological (Moses and Gonzales 2015) and the cognitive-linguistic (Wolters 2012) annotation schemes.

	Results social-psychological method		Results cognitive-linguistic method	
	Strict Father	Nurturant Parent	Strict Father	Nurturant Parent
Complete corpus	OR (BO) = 1.12 $\chi^2(1) = 2.70$ , $p = 0.10$ , Bush = Obama	OR (BO) = 1.36 $\chi^2(1) = 17.54$ , $p < 0.05$ , Bush < Obama	OR (GB) = 1.03 $\chi^2(1) = 0.04$ , $p = 0.85$ , Bush = Obama	OR (BO) = 1.12 $\chi^2(1) = 0.65$ , $p = 0.42$ , Bush = Obama
Speeches Iraq War	OR (BO) = 1.22 $\chi^2(1) = 4.85$ , $p < 0.05$ , Bush < Obama	OR (BO) = 1.36 $\chi^2(1) = 8.23$ , $p < 0.05$ , Bush < Obama	OR (BO) = 1.11 $\chi^2(1) = 0.40$ , $p = 0.53$ , Bush = Obama	OR (BO) = 1.38 $\chi^2(1) = 3.50$ , $p = 0.06$ , Bush = Obama
Speeches Health Care	OR (BO) = 1.22 $\chi^2(1) = 2.34$ , $p = 0.13$ , Bush = Obama	OR (BO) = 1.36 $\chi^2(1) = 9.09$ , $p < 0.05$ , Bush < Obama	OR (GB) = 1.22 $\chi^2(1) = 0.41$ , $p = 0.52$ , Bush = Obama	OR (GB) = 2.87 $\chi^2(1) = 14.52$ , $p < 0.05$ , Bush > Obama

Notes: OR (GB) refers to the odds ratio (OR) of Bush using a certain expression over Obama, and OR (BO) refers to the odds ratio of Obama using a certain expression over Bush. Thus, for example, an OR (GB) of 2.00 indicates that it is 2.00 times as likely that Bush uses a certain expression over Obama. An OR (BO) of 2.00 indicates that is 2.00 times as likely that Obama uses a certain expression over Bush.

applied to the same corpus, and whether they could thus be considered to be equally appropriate ways of testing this theory. This is relevant since both annotation schemes claim to measure the same phenomenon, and they should therefore yield comparable results with regard to whether Bush and Obama use similar or different amounts of Strict Father or Nurturant Parent language. The social-psychological approach to data analysis identified any expression that could be considered an example of one of the two models, without considering the metaphoricality of an expression (Mosses and Gonzales 2015). In contrast, the cognitive-linguistic annotation scheme focused on finding metaphorical expressions that could be ascribed to one of the two political moral models (Wolters 2012). The two annotation schemes were applied to the same corpus, which comprised twenty speeches. Ten of these speeches were delivered by George W. Bush, and the other ten by Barack Obama. Half of the speeches were on the topic of Health Care, the other half on the war in Iraq.

This study shows that conclusions based on methods such as discussed in this paper should be treated with caution, because the results from this study indicate that the two distinct methods of data coding under study generated diverging results, which implies that the two methods are thus not equally appropriate ways of identifying moral language in political discourse. We find a number of differences and complications concerning the two methods, which are possible explanations for the diverging results.

The first difference is that Moses and Gonzales (2015) identified any expression that could be explained by the Strict Father or Nurturant Parent models without considering metaphoricity of the expression, while Wolters (2012) only identified metaphorical expressions reflecting the conceptual metaphors for morality belonging to either the Strict Father or Nurturant Parent model. This would not necessarily be problematic if non-metaphorical language would in fact reflect the logic of the two politically moral models in a comparable way to metaphorical language (Cienki 2005). However, since the results of the two annotation schemes are diverging, this does not seem to be the case. Since the cognitive models as described by Lakoff (2002 [1996]) are metaphorical by nature, and CMT posits that conceptual metaphors are reflected by metaphorical expressions in language, studies aimed at identifying moral language in political discourse following the hypotheses of TMR should focus on metaphorical language, and not expand beyond these.

However, it is surprising how many discourse units in our corpus are moral statements, while the number of relevant metaphorical lexical items that we found is very small. This might be due to the fact that Wolters (2012) only considered a limited number of lexical items, and we therefore did not identify all possibly relevant metaphorical Strict Father or Nurturant Parent lexical items in our corpus. It could thus be that in the discourse units that were identified as either Strict Father or Nurturant Parent language by the social-psychological approach, other relevant metaphorical items were present, but that they were not identified since Wolters did not take them up in his list of relevant lexical items. A check of the data reveals that there are indeed a number of other lexical items that could be relevant to TMR. A few examples of metaphorically used lexical items that could be relevant for the Strict Father morality model that were found in the speeches are *relief* (when referring to taxes), *to guard*, *firm*, *straight*, and *healthy*. There were also some metaphorically lexical items found that could reflect the Nurturant Parent model, such as *to stand*, *to lend (a hand)*, and *to bear*. Nevertheless, it remains questionable whether there are many more of these examples, especially since Cienki (2005) also only found a small number of metaphorical expressions directly reflecting the central Strict Father and Nurturant Parent metaphors.

The paucity of metaphorical lexical items referring to morality raises the question whether there is not a better way to explain the moral differences between conservatives and liberals than through a metaphorical model. In order to answer this question, future research into TMR should not limit itself to a selection of possibly relevant lexical units, but identify all possibly relevant metaphorical expressions. This way, a better estimation of the number of relevant metaphorical expressions can be obtained.

The second difference between the two annotation schemes concerns the categories that the coding taxonomies of the two studies comprised. Lakoff (2002 [1996]) considered eight prioritized metaphors for morality for each of the two family models. The eight categories belonging to the Strict Father model are Moral Strength, Moral Authority, Moral Essence, Moral Wholeness, Moral Order, Moral boundaries, Moral Purity, and Moral Health. The ones belonging to the Nurturant Parent model are Morality as Empathy, Morality as Nurturance, Moral Self-Nurturance, Morality as Social Nurturance, Morality as Happiness, Morality as Self-development, Morality as Fair Distribution, and Moral Growth. The taxonomy of Moses and Gonzales (2015) comprised a number of these prioritized metaphors for morality and a number of subcategories, resulting in eight Strict Father categories and six Nurturant Parent categories. Contrarily, Wolters' (2012) coding taxonomy consisted of the eight prioritized metaphors for morality for each of the two models as described by Lakoff (2002 [1996]). The two methods only correspond on three categories, namely the two Strict Father categories Moral Strength and Moral Authority, and the Nurturant Parent category Morality as Nurturance. Moses and Gonzales (2015) thus left out six of the superordinate categories belonging to the Strict Father model, and seven belonging to the Nurturant Parent model.

The fact that Moses and Gonzales (2015) did not consider all eight prioritized categories as described by Lakoff (2002 [1996]) might have influenced the number of expressions that were identified to be reflections of the Strict Father and Nurturant Parent models. Various expressions have perhaps not been coded as belonging to one of the two models because they belong to a category that is not represented in Moses and Gonzales' (2015) taxonomy. This could also have had its' influence on the results found with regard to the comparability of the two presidents and their use of Strict Father and Nurturant Parent language, and to the comparability of the two annotation schemes. A preliminary check of the data shows that there are indeed a number of discourse units that have not been coded as belonging to either the Strict Father or Nurturant Parent model, while they are in fact reflections of one of the categories that were not included by Moses and Gonzales (2015). The utterance *There ought to be transparency in medicine* for example, is an instance of the Strict Father category Moral Purity,



because the word *transparency* is part of Wolters' (2012) list of relevant lexical items for this Strict Father category. Another example is the utterance *And that way, you begin to make sure the tax code is a level playing-field*, which is an instance of the Nurturant Parent category Morality as Fair Distribution. The online Macmillan dictionary gives the following description for the word *level playing-field*: "A situation that is fair for all the people involved". With this utterance the speaker shows he feels that the tax code should be fair for all American citizens. The annotation scheme as presented by Moses and Gonzales (2015) thus does not identify all relevant discourse units referring to the Strict Father or Nurturant Parent models. Their annotation scheme is selective in nature, since they only considered a selection of the categories relevant to the Strict Father and Nurturant Parent models as presented by Lakoff (2002 [1996]). In order to be able to identify all relevant discourse units, future research should include all categories as described by Lakoff (2002 [1996]).

Besides these main differences between the two annotation schemes, there is another issue to consider, and that is the lack of clear guidelines in their taxonomies. The combination of superordinate categories with the subcategories in the social-psychological approach made it difficult to distinguish between a number of categories such as Morality as Strength and Morality as Self-Discipline or Morality as Nurturance and Responsibility for Others, especially since the categories were not very well defined or made distinct from each other. In addition, both the social-psychological and the cognitive-linguistic annotation scheme did not explicitly define and operationalize morality or what a moral expression constitutes. In trying to solve this issue, we added the steps of deciding whether a discourse unit was referring to politics and morality, and subsequently, whether it could be coded for belonging to the Strict Father or Nurturant Parent model. Since Moses and Gonzales (2015), and Wolters (2012) did not define or operationalize the concept morality in their studies, it is impossible to say how our definition and operationalization differs from theirs, and how this consequently might have led to different results. With regard to transparency of methodology, consistency of coding, and replicability of research, we suggest that future research into TMR gives a definition and operationalization of key concepts such as morality.

When we consider the results yielded by the two distinct annotation schemes, another observation can be made: both the results from the social-psychological and the cognitive-linguistic annotation scheme do not offer strong support for Lakoff's assumptions with regard to moral reasoning in American politics. The results of the social-psychological annotation scheme do indicate that Obama used more Nurturant Parent language than Bush. However, based on the odds ratio, the odds of Obama using a Nurturant

Parent expression is only 1.36 times larger than Bush. Furthermore, the results of this annotation scheme indicate that Bush does not use more Strict Father language than Obama. Within the speeches on the war in Iraq, the odds of Obama using a Strict Father expression was even 1.22 times larger than Bush. The results of the cognitive-linguistic annotation scheme suggest that political party (Republican vs Democratic President) has no impact on the President's use of Strict Father or Nurturant Parent language. It could, however, be that the genre of political speeches does not lend itself to too strong an emphasis on one of the two models, because a politician needs to appeal to a broader, bi-partisan audience, explaining why we did not find strong evidence for Lakoff's assertions on moral reasoning in American politics. Another possibility is that the level of Strict Father and Nurturant Parent language used by conservatives and liberals depends on the topic that is addressed. In order for future research to be able to answer these questions, it should examine corpora treating several different political topics. Future research can also focus on discussions by people who are not in politics, and who are thus not restrained by having to appeal to a broad audience, but do identify with either conservatives or liberals.

In our statistical analysis, we took both effect size (i. e. odds ratios) and statistical significance (i. e. *p*-values) into account. In doing so, we follow the suggestion of Kotrlik and Williams (2003), who state that reporting on statistical significance alone is not enough, and that to be able to judge the magnitude of the differences present between groups, effect sizes are also necessary. Statistical testing is an issue that is barely addressed by cognitive-linguistic scholars, and most studies seem to focus on *p*-values only when reporting results. For example, both Wolters (2012) and Cienki (2005) did not report on effect sizes. By reporting on both effect sizes and statistical significance, we show whether a statistically significant outcome is or is not practically meaningful, and whether a practically meaningful outcome did or did not occur by chance (Kotrlik and Williams 2003).

A last issue that should be addressed is the fact that, in order to be able to compare the two annotation schemes, and to raise their reliability, a number of adjustments were made to the original methods as presented by Moses and Gonzales (2015), and Wolters (2012). The methods as used in this study thus slightly differ from the original methods.

To conclude, this study demonstrates that conclusions based on results obtained by using annotation schemes such as discussed in this paper should be taken with caution, because choosing one method over the other can possibly be a large factor in the eventual results. We thus recommend that when testing TMR, the corpus linguistic method is critically evaluated.

**Acknowledgment:** All complete coding schemes and data files are fully available from the Open Science Framework (OSF) at <https://osf.io/v89nq/>.

When the research was conducted, Kiki Y. Renardel de Lavalette was a Master Student in Communication and Information Sciences at Vrije Universiteit Amsterdam (the Netherlands). She is now a PhD candidate at the Department of Dutch Studies at the University of Amsterdam, the Netherlands. The authors would like to thank Dr. Stefanie Wulff and two anonymous reviewers for their helpful comments and suggestions.

**Funding:** The contribution of Christian Burgers was supported by the Netherlands Organisation for Scientific Research (NWO VENI grant 275-89-020).

## References

- Bar-Lev, Zev. 2007. Reframing moral politics. *Journal of Language and Politics* 6(3). 459–474. doi: 10.1075/jlp.6.3.11bar
- Charteris-Black, Jonathan & Andreas Musolff. 2003. ‘Battered hero’ or ‘innocent victim’? A comparative study of metaphors for euro trading in British and German financial reporting. *English for Specific Purposes* 22. 153–176. doi: 10.1016/S0889-4906(02)00012-1.
- Cienki, Alan. 2005. Metaphor in the “Strict Father” and “Nurturant Parent” cognitive models: Theoretical issues raised in an empirical study. *Cognitive Linguistics* 16(2). 279–312. doi: 10.1515/cogl.2005.16.2.279
- Deason, Grace & Mari H. Gonzales. 2012. Moral politics in the 2008 presidential convention acceptance speeches. *Basic and Applied Social Psychology* 34(3). 254–268. doi: 10.1080/01973533.2012.674450.
- Deason, Grace, Brad Lippmann, Marti H. Gonzales & Jennifer Filson. 2008. Exploring Uncle Sam: The role of family metaphors in political attitudes. Paper presented at the annual meeting of the International Society of Political Psychology, Paris, France, 08 July.
- Koller, Veronika. 2008. *Metaphor and gender in business media discourse. A critical cognitive study*. New York: Palgrave Macmillan.
- Kotrlik, Joe W. & Heather A. Williams. 2003. The incorporation of effect size in information technology, learning, and performance research. *Information Technology, Learning, and Performance Journal* 21(1). 1–7.
- Lakoff, George. 2002 [1996]. *Moral politics. How liberals and conservative think*. Chicago: The University of Chicago Press.
- Lakoff, George & Mark Johnson. 1980. *Metaphors we live by*. Chicago: University of Chicago Press.
- Landis, J. Richard & Gary G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33(1). 159–174. doi: 10.2307/2529310
- Level playing-field. n.d. In *MacMillan Dictionary online*. Retrieved from <http://www.macmillan.com/dictionary/british/level-playing-field>.

- Moses, Jennifer F. & Marti H. Gonzales. 2015. Strong candidate, nurturant candidate: Moral language in presidential television advertisements. *Political Psychology* 36(4). 379–397. doi: 10.1111/pops.12160.
- Ohl, Jessy J., Damien S. Pfister, Martin Nader & Dana Griffin. 2013. Lakoff's Theory of Moral Reasoning in presidential campaign advertisements, 1952–2012. *Communication Studies* 64(5). 488–507. doi: 10.1080/10510974.2013.832340.
- Pragglejaz Group. 2007. MIP: A method for identifying metaphorically used words in discourse. *Metaphor and Symbol* 22(1). 1–39. doi: 10.1080/10926480709336752
- Semino, Elena. 2008. *Metaphor in discourse*. Cambridge: Cambridge University Press.
- Steen, Gerard J. 2005. Basic discourse acts: Towards a psychological theory of discourse segmentation. In Francisco J. Ruiz de Mendoza Ibáñez & M. Sandra Peña Cervel (eds.), *Cognitive linguistics: Internal dynamics and interdisciplinary interaction*, 283–312. Berlin & New York: Mouton de Gruyter.
- Steen, Gerard J., Aletta G. Dorst, J. Berenike Herrmann, Anna Kaal, Tina Krennmayr & Trijntje Pasma. 2010. *A method for linguistic metaphor identification: From MIP to MIPVU*. Amsterdam: John Benjamins.
- Wolters, H.C. 2012. *Is Bush a strict father and Obama a nurturing parent? Metaphorical expressions of moral value systems in American politics*. Amsterdam: Vrije Universiteit Amsterdam MA thesis.