Appendix A

Attitude issue descriptions

Migration

"People migrate for a variety of reasons, including, to study, to work, to reunite with family, or to escape violence and persecution in their home country. There is currently a migration crisis worldwide, however, host countries have limitations with respect to how many migrants they can accommodate. Decisions with respect to migrant intake are detailed in migration policy. Recently, governments and political parties have expressed different views on migrant intake quotas".

Carbon emissions

"Carbon dioxide emissions from electricity and gas usage contribute to more than half of Victoria's total carbon dioxide emissions. These emissions peak particularly during colder times of the year, due to significantly increased usage of heating. The higher the household thermostat temperature, the greater the energy consumed. The Australian government may therefore pass a law specifying that households set a maximum thermostat temperature of 20 degrees Celsius during winter."

Appendix B

Advocacy Task Instructions

Advocacy task (Moral). While our focus was tapping into the five moral foundations (harm, fairness, loyalty, authority, purity) (see Haidt, 2007; Haidt, 2001), instructions for the moral condition were created based on both moral psychology and positive psychology, to capture the broad range of values and virtues which tap into morality (see Peterson & Seligman, 2004 for moral virtues). This was in order to ensure that participants were not constrained in the expression of their attitudes. In the moral advocacy task, participants were

encouraged to think about a moral choice as "weighing issues of harm, rights, justice and fairness", considering rights violations, and making decisions "with respect to a set of virtues in a culture or subculture" (Haidt, 2001).To capture these concerns, our index of moral expressiveness was scored on the binding moral foundations (average scores on loyalty, authority, purity) and individualizing moral foundations (average scores on harm and fairness; see Jonathan Haidt & Joseph, 2004, 2008). Participants were presented with the following instructions:

"People use different kinds of arguments to persuade others. For example, they may use arguments grounded in moral values. On the other hand, they may use arguments grounded in practical concerns, such as, resource availability. For this task, we ask you to focus on <u>fundamental moral values </u>which relate to <u>your position </u>on migration. We are <u>not</u> interested in practical concerns related to your position. Imagine that you have to persuade someone who disagrees with you on your view about how many migrants the American government should take**.** Convince this person to adopt your <u>own </u>position on this issue. Specifically, write arguments **only** from a **moral** perspective. That is, base your arguments on moral considerations such as (but not restricted to) the following:

- Whether something is fundamentally right or wrong, good or bad.

- Whether something is worthy of blame or punishment or moral praise.

- Moral values relating to concerns such as care, harm, injustice, fairness, loyalty, betrayal, disrespect, purity or liberty.

- Virtues such as wisdom, courage, generosity, transcendence, justice, prudence".

Advocacy task (Practical). Instructions for the practical condition were created based on economic theory, to capture concerns related to resource scarcity. In the practical advocacy task, participants are encouraged to think about a practical or rational choice as one that "maximizes expected utility" (Kahneman & Tversky, 1986; Tversky & Kahneman,

1981). It is thinking in terms of outcomes (e.g. if X happens then Z consequences), in terms of money, and resources, for example. Participants were told that a practical choice is on which allocates scarce resources to the alternative which maximizes positive outcomes for the group; "the rational man is a maximiser" (Simon, 1978). It is balancing consequences (i.e. cost-benefit analysis) to pursue goals which "maintain the integration and adaptation of the larger system" (Simon, 1978).

"People use different kinds of arguments to persuade others. For example, they may use arguments grounded in moral values. On the other hand, they may use arguments grounded in practical consequences, such as, resource availability.

For this task, we ask you to focus on practical concerns which relate to your position on migration. We are <u>not</u> interested in moral values related to your position.

Imagine that you have to persuade someone who disagrees with you on your view about how many migrants the American government should take. Convince this person to adopt your <u>own</u> position on this issue.

Specifically, write arguments **only** from a **practical** perspective. That is, base your arguments on practical considerations such as (but not restricted to) the following:

- Whether something is feasible or workable, pragmatic or realistic

- Whether something is based on evidence, on something concrete and factual

- Whether something is grounded in rational cost-benefit analyses (e.g. costs outweighing benefits, or benefits outweighing costs)

 - Consider things like resource considerations (e.g., how resources are allocated), economic consequences."

Appendix C

Constructing Dictionaries

In constructing the dictionaries, we generally aimed to avoid words with common but irrelevant alternate meanings. For example, the word *right* (in the sense of "human right") is relevant to deontological ethics, however the word is arguably more commonly used to reference either correctness (e.g., "the right answer") or location (e.g., "to the right"). Because some dictionary categories were represented by a very small number of terms, we performed a wildcard search through the model vocabulary to identify additional variants of our keywords. This strategy serves the goal of boosting the reliability of the dictionary by including vectors for relevant variants of a word (much like adding additional items to a questionnaire). Additionally, by including variants of the same word we are able (to some extent) to average out idiosyncratic or unwanted alternate meanings of some of the words. For example, in the case of the dictionary intended to capture expressions of *Contempt*, by including multiple variants of the word *contempt* (e.g., *contemptible*), the common use of *contempt* as a legal term (as in "contempt of court," which doesn't reference any actual emotional state) should have relatively less influence on the concept representation.

To identify the words that would define each concept, we went through the Wheeler and Laham (2016) dictionary, and identified the most important words for each dictionary category. For each of these words, we performed a wildcard search through the model vocabulary so we could average over relevant variants in the dictionary (e.g., averaging over variants of **anger**, **rage**, etc.) before averaging over the composite vectors to derive our representation of Anger. The rationale for the two-step averaging (within word stems and *then* across word stems) rather than just averaging all terms together was that different words had vastly different numbers of variants (e.g., there were two relevant variants of *furious*, but five relevant variants of *anger*), however we didn't want the category vectors to be

substantially weighted in favor of whatever words happen to have the most variants, because it's unlikely that the number of variants of a word corresponds to their importance for our dictionary categories. After generating the wildcard list, it was manually inspected by the authors who decided which words to include in the final dictionary.

For the emotion dictionary concepts, we decided to omit antonyms, because although such terms are obviously semantically related, their inclusion would be inconsistent with the goal of creating dictionaries that capture the *expression* of the particular emotion. For the non-emotion dictionaries however, we decided to include antonyms (e.g., having *responsibility* and *irresponsibility* both representing *Deontology*) because both arguably express consideration of that particular moral code.

The full list of words used to capture all dictionary concepts are indicated in the table below.

Concept	Words
Moral foundations	
Harm	suffer
	cruel
	hurt
	harm
Care	kindness
	compassion
	nurture
	empathy
Fairness	fairness
	equality

	justice
	rights
Cheating	cheat
	fraud
	unfair
	injustice
Subversion	subversion
	disobey
	disrespect
	chaos
Loyalty	loyal
	solidarity
	patriot
	fidelity
Betrayal	betray
	treason
	disloyal
	traitor
Authority	authority
	obey
	respect
	tradition
Purity	purity
	sanctity
	sacred

	wholesome
Degradation	impurity
	depravity
	degradation
	unnatural
Moral systems	
Deontology	duty
	obligation
	obligations
	obligated
	obligatory
	obligate
	obligates
	obligating
	rights
	righteous
	righteousness
	rightful
	rightfully
	righteously
	human-rights
	civil-rights
	rights-based
	responsible
	responsibility

	responsibilities
	irresponsible
	responsibly
	irresponsibility
	irresponsibly
	responsiblity
	responsibilty
	responsibilites
	responsiblities
	rules
	rule
	rule-breaking
	norm
	norms
	normative
Consequentialism	outcome
	outcomes
	outcome-based
	consequences
	consequence
	consequential
	consequent
	consequentially
	consequentialism
	consequentialist

repercussion

repercussions

repercussion

cost

costs

costly

cost-effective

costing

low-cost

cost-effectiveness

cost-effectively

cost-benefit

cost-efficient

costlier

lower-cost

high-cost

cost/benefit

costliest

cost-based

cost-efficiency

lowest-cost

benefits

benefit

beneficial

benefited

	beneficiaries
	beneficiary
	benefiting
	benefitted
	benefitting
	beneficent
	beneficially
	beneficiating
	beneficence
	risk-benefit
	benefic
	benefit-cost

Moral emotions

Anger	angry
	anger
	angered
	angers
	angering
	outrageous
	outrage
	outraged
	outrageously
	outrages
	outrageousness
	furious

	furiously
	rage
	enraged
	rages
	raged
	enrage
	frustrated
	frustrating
	frustration
	frustrations
	frustrate
	frustrates
	frustratingly
Contempt	contempt
	contemptuous
	contemptible
	contemptuously
	despise
	despised
	despises
	despising
	detest
	detested
	detestable
	detests

	detestation
Disgust	disgusting
	disgust
	disgusted
	disgustingly
	disgusts
	repulsive
	repulsed
	repulsion
	repulse
	revolting
	sickening
	sickened
	sickens
	sicken
	sickeningly