

1 Fixed versus growing seed words

Here, we compare three situations:

- **fixed seeds:** throughout the analysis, only access the seed words available in 1800;
- **varying seeds:** in each decade, evaluate the model on all seed words available then;
- **restricted seeds:** in each decade, use all seed words available in that decade to make predictions, *but only measure predictive accuracy over the seed words available in 1800*.

The reason to look at the third, unorthodox-looking option is to distinguish between two possible reasons for degraded performance in more recent decades: 1) more recently-available words are harder, and 2) more recently-available words add noise to the model and affect previously-correct predictions.

Figures 1 to 3 suggest that words whose embeddings are only available more recently are, in fact, more difficult to classify.

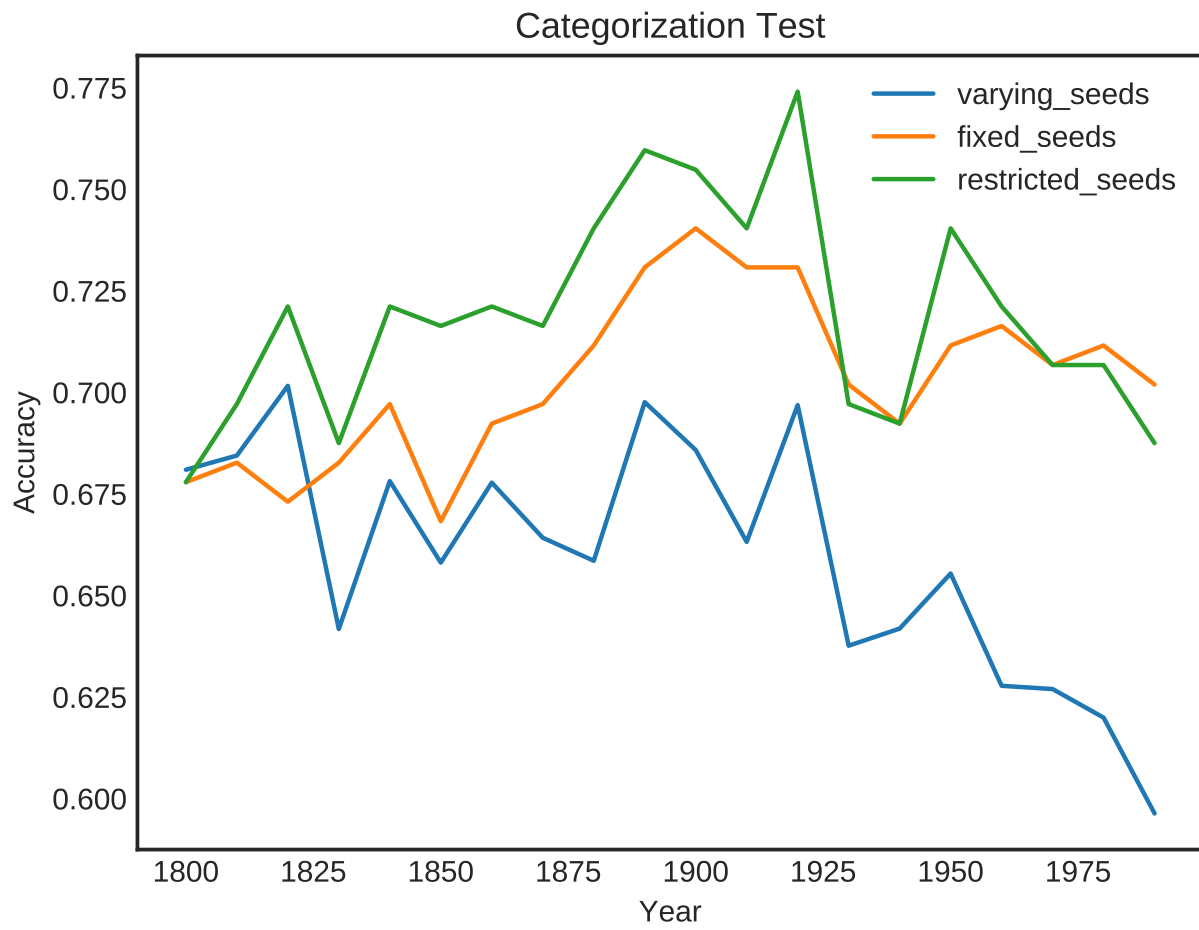


Figure 1: Categorization accuracy for different seed word regimes

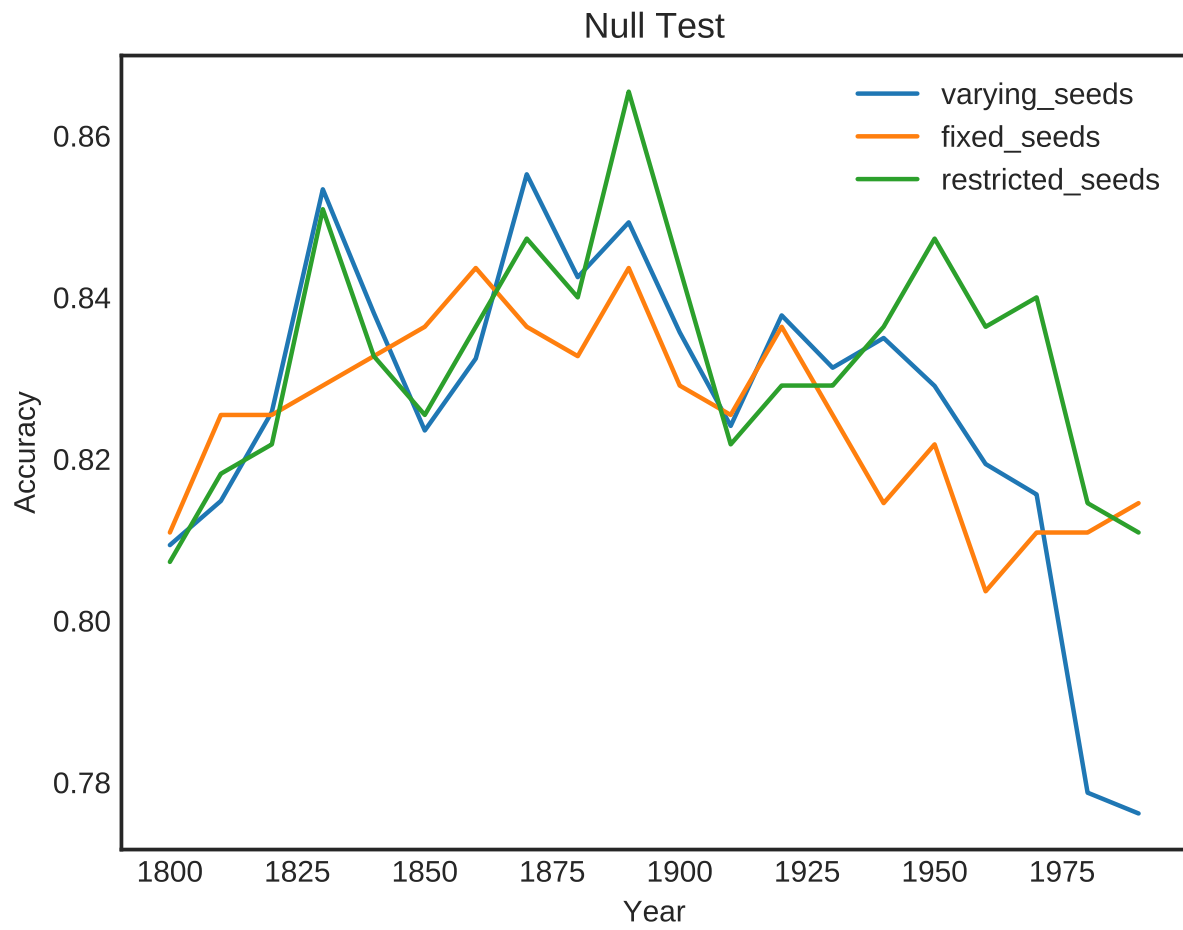


Figure 2: Null test accuracy for different seed word regimes

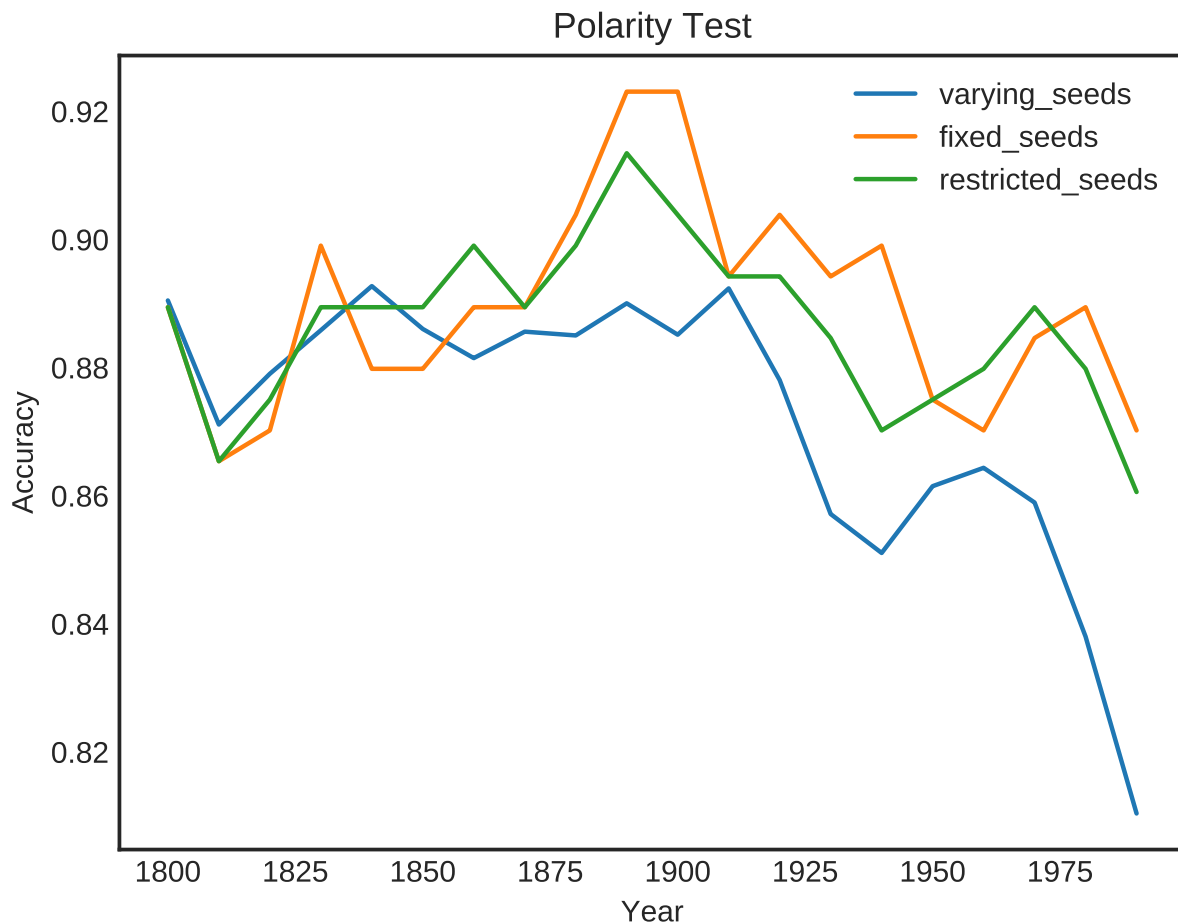


Figure 3: Polarity test accuracy for different seed word regimes

2 FDA + centroid classifier

We repeat the classification tests using the centroid method, but pre-process word embeddings in each decade by applying FDA dimension reduction to them.

More precisely, before predicting each seed word s in the leave-one-out procedure, we apply FDA to the set of all seed words *except for* s , and use the coefficients obtained to project s into the same space. Then, we classify s as before in the reduced space.

Since we now have 11 categories in total (10 moral categories along with a

neutral category), we perform FDA over 10 dimensions (instead of 9). We apply shrinkage coefficient 0.5, and use the `eigen` solver of `scikit-learn`.

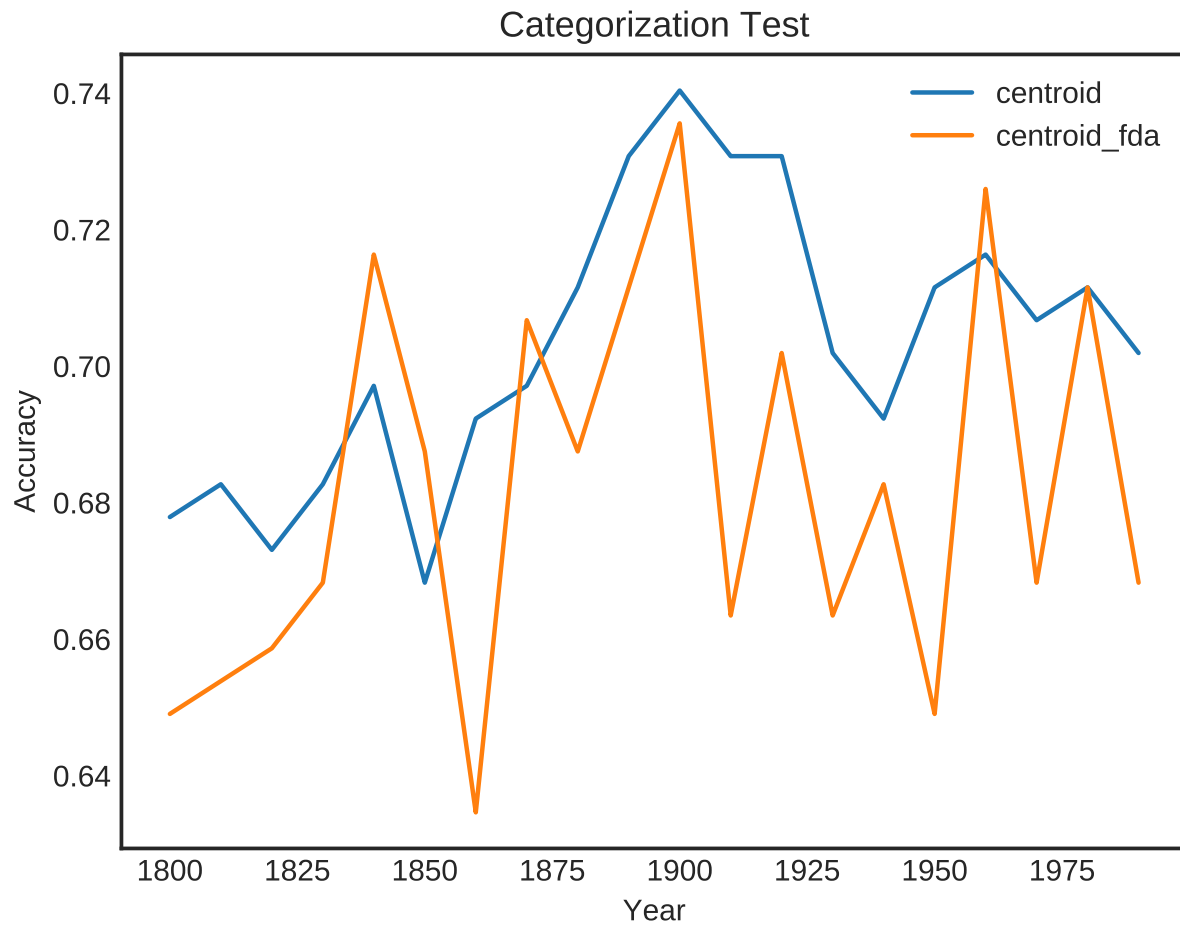


Figure 4: Categorization accuracy with centroid and FDA+centroid methods

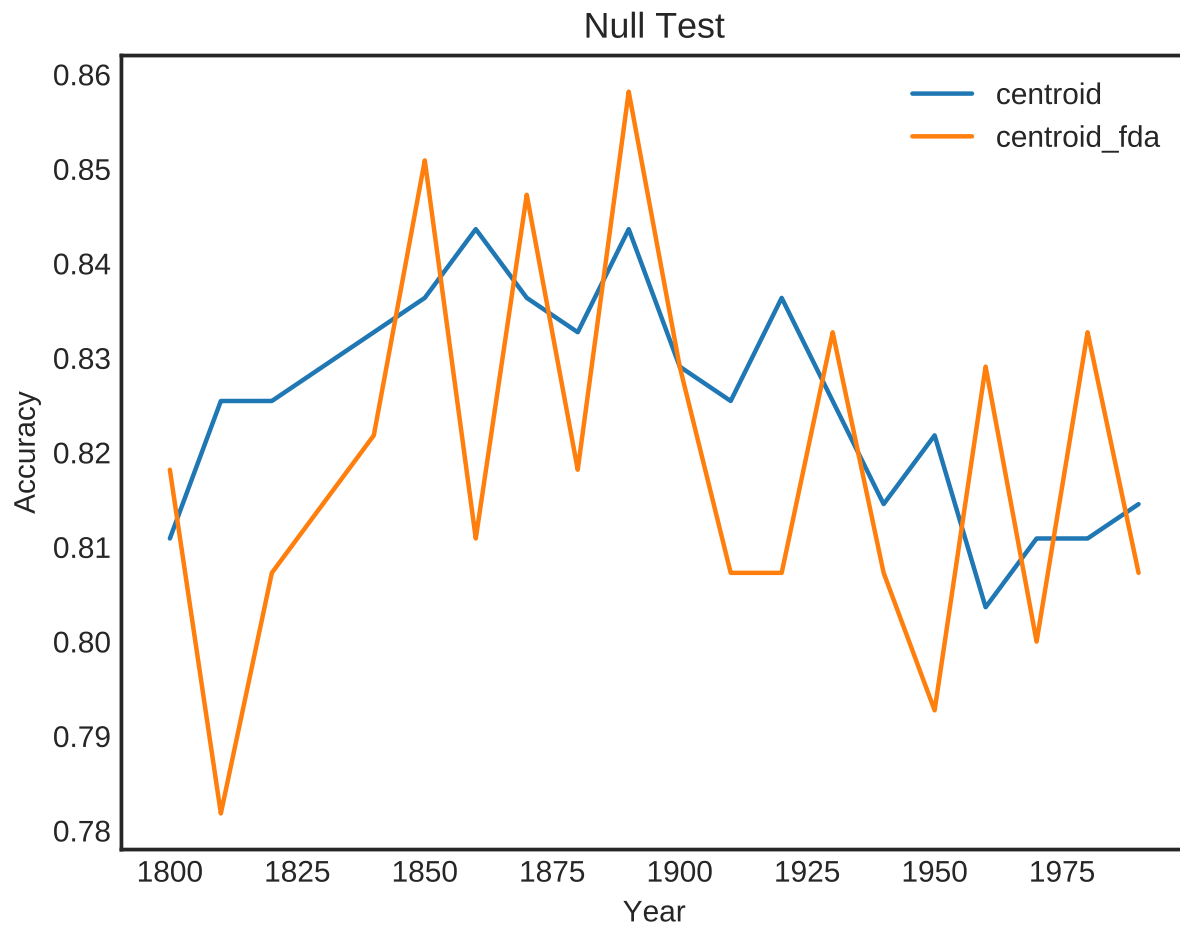


Figure 5: Null test accuracy with centroid and FDA+centroid methods

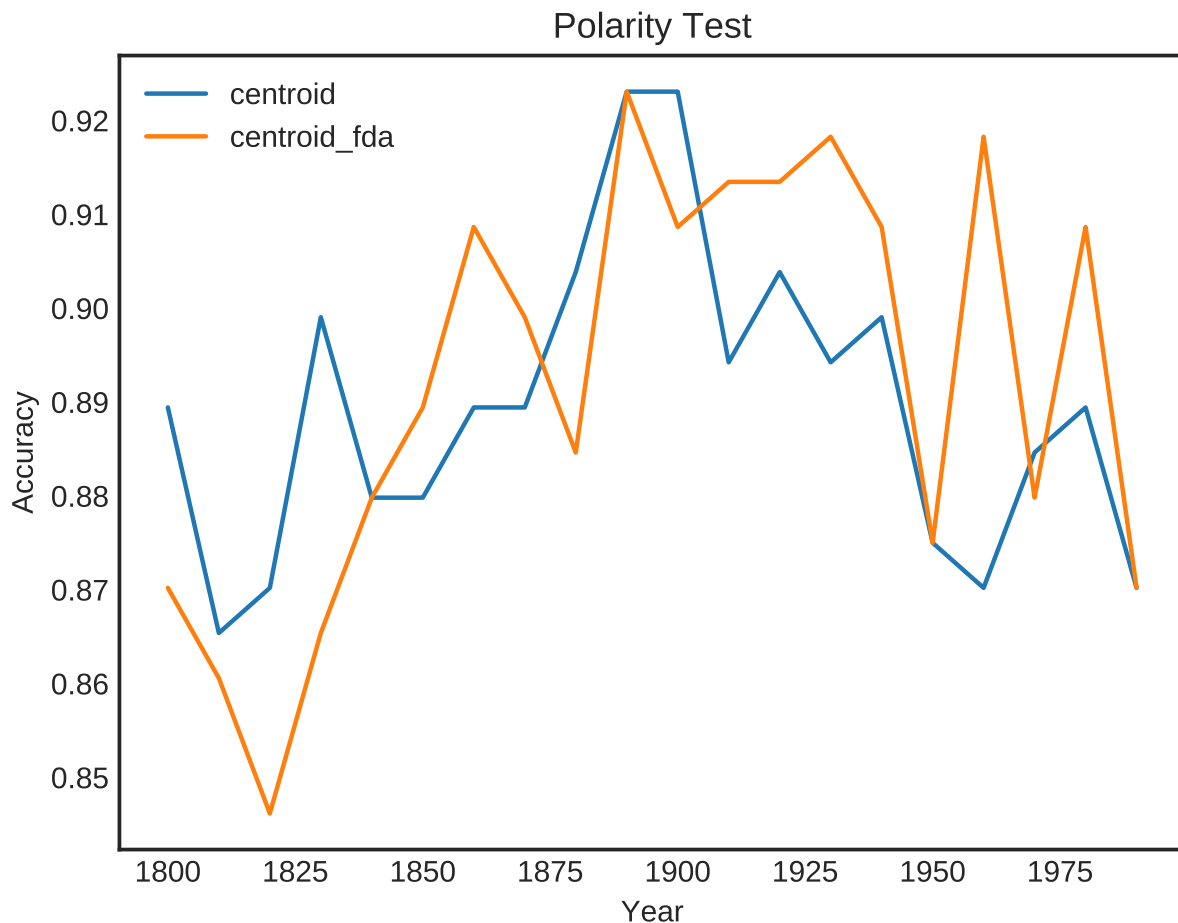


Figure 6: Polarity test accuracy with centroid and FDA+centroid methods

Figures 4 to 6 show classification accuracy in the **fixed seeds** regime (same seed words as 1800) for both methods. Applying FDA prior to classification does not seem to systematically improve predictive accuracy.

3 Confusion matrices

Figures 7 to 12 show confusion matrices for each classification test under each of the fixed seeds and varying seeds regimes, both of which display similar patterns.

The category **-loyalty** seems particularly challenging, and seems to be confused with **-sanctity**. The number of available seed words for each category

may play a contributing role (see Table 1), but it is not yet clear how significant this factor is versus other potential causes.

NOTE: One possible idea is to artificially rebalance classes, at least somewhat, and see whether this anomaly improves. I (RF) suspect not, and that the problem is semantic in nature.

NOTE: Once kernel density procedures have finished, we will be able to see how that model compares in this particular challenging case.

TODO: Consider enlarging neutral seed word set to match positive + negative words, and whether to do so pre- or post- embedding filtering (bias?).

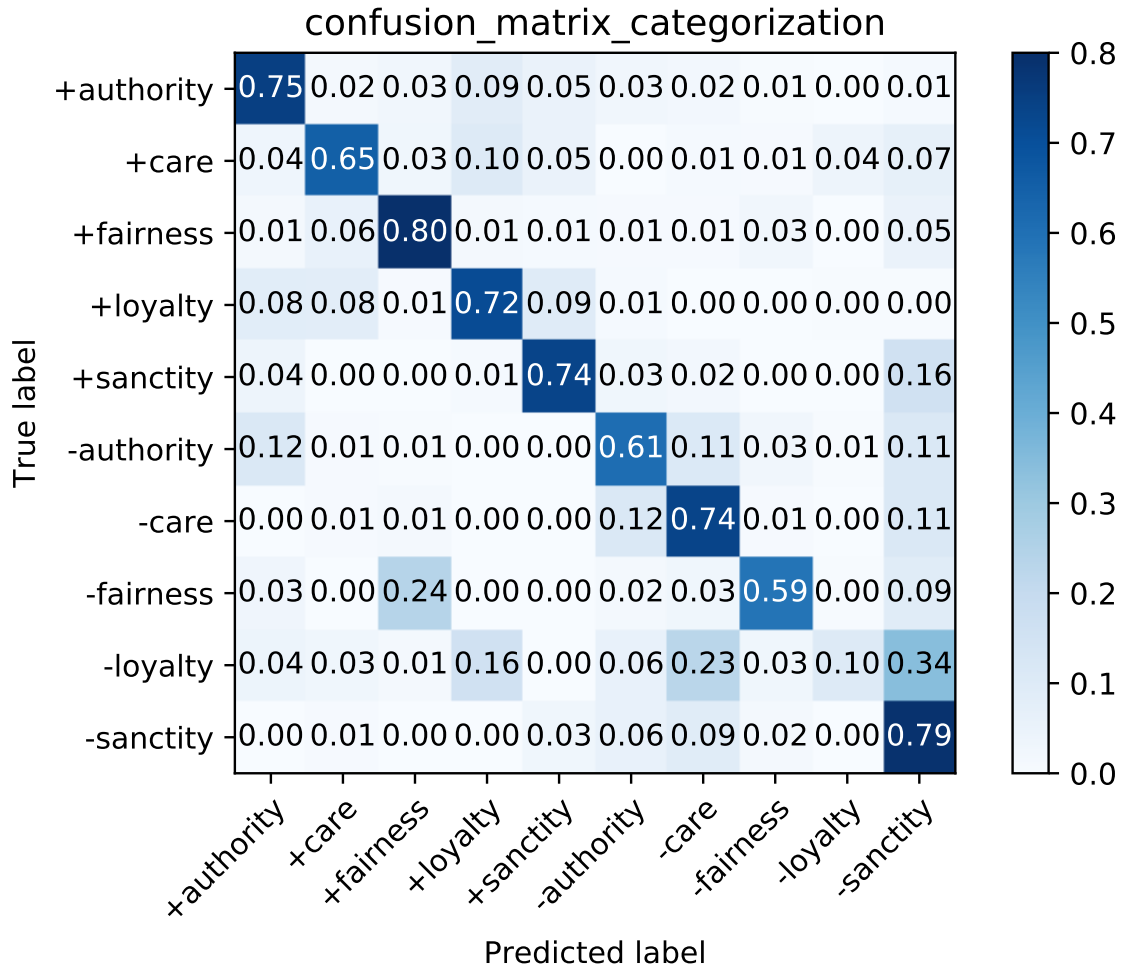


Figure 7: Categorization confusion matrix for fixed seeds.

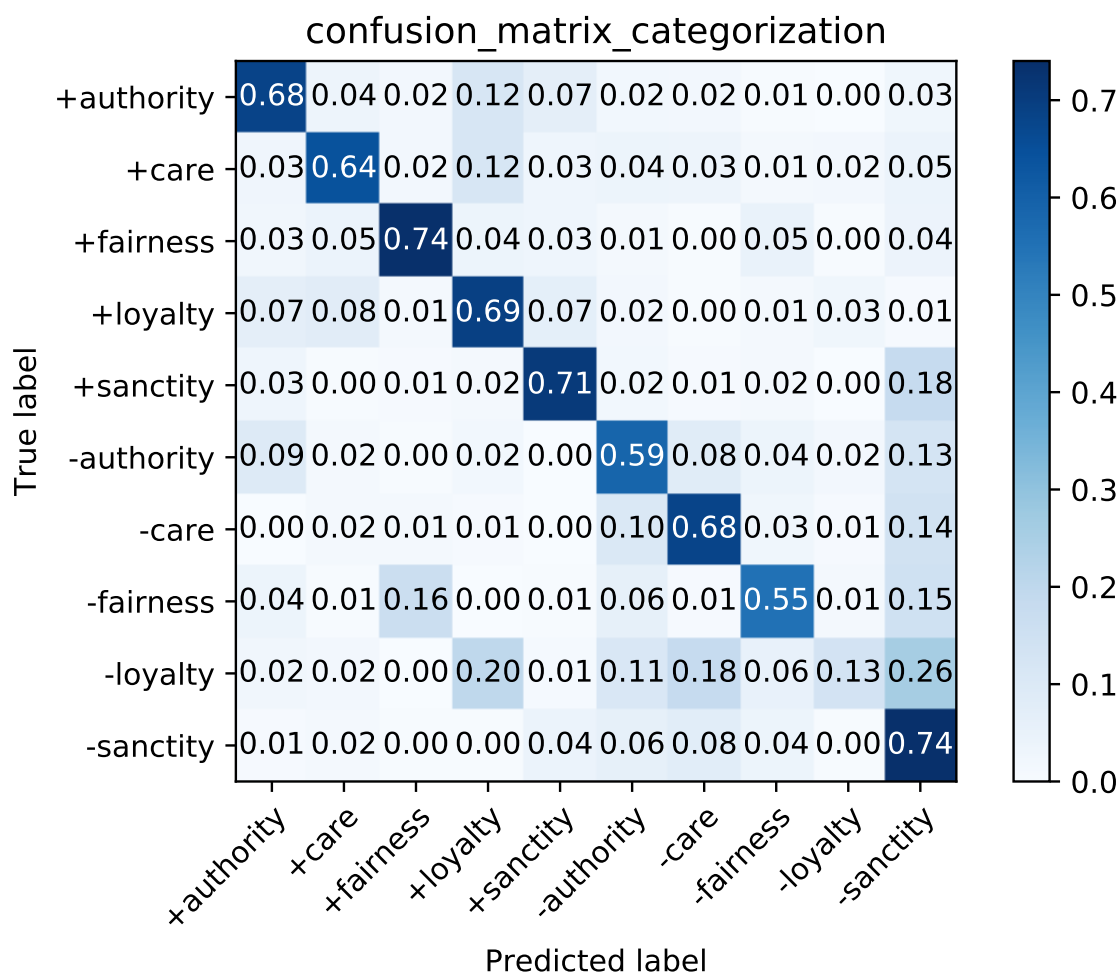


Figure 8: Categorization confusion matrix for varying seeds.

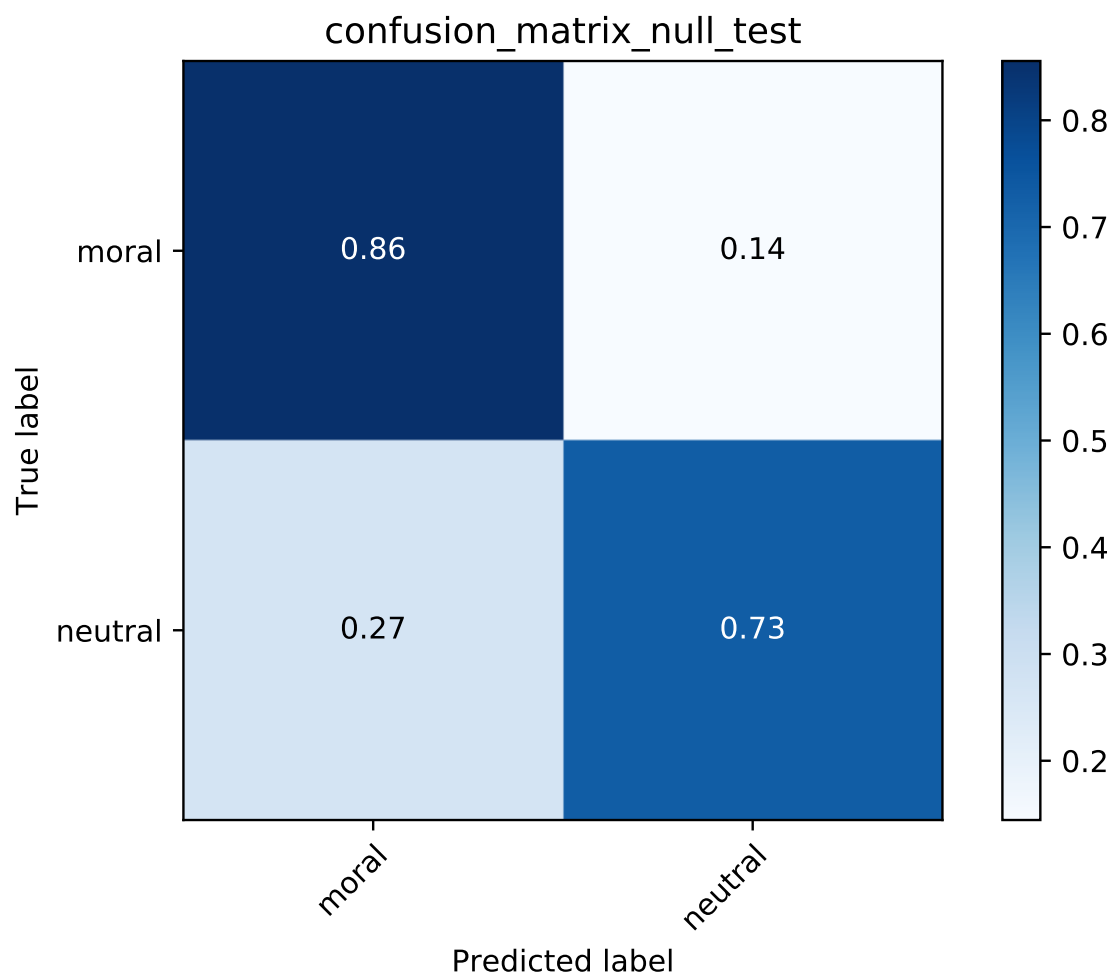


Figure 9: Null test confusion matrix for fixed seeds.

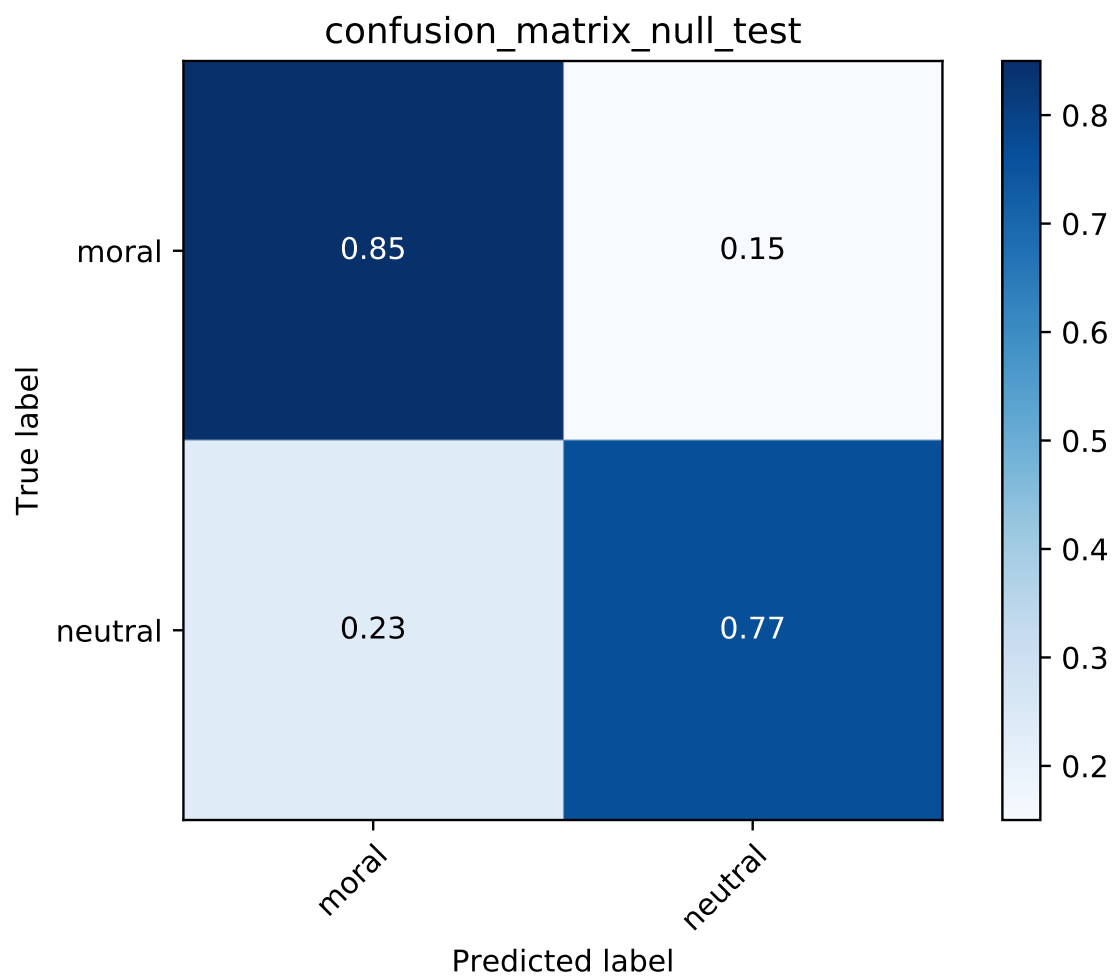


Figure 10: Null test confusion matrix for varying seeds.

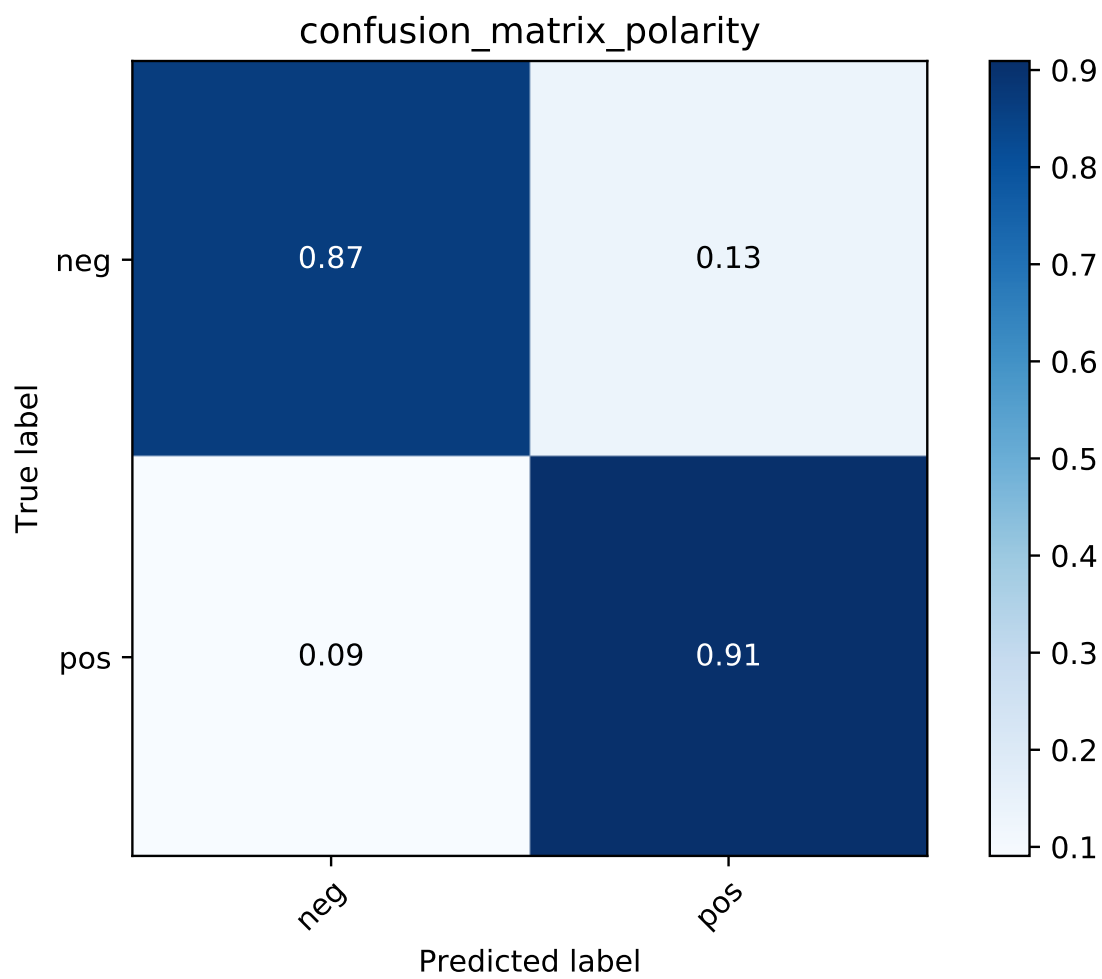


Figure 11: Polarity test confusion matrix for fixed seeds.

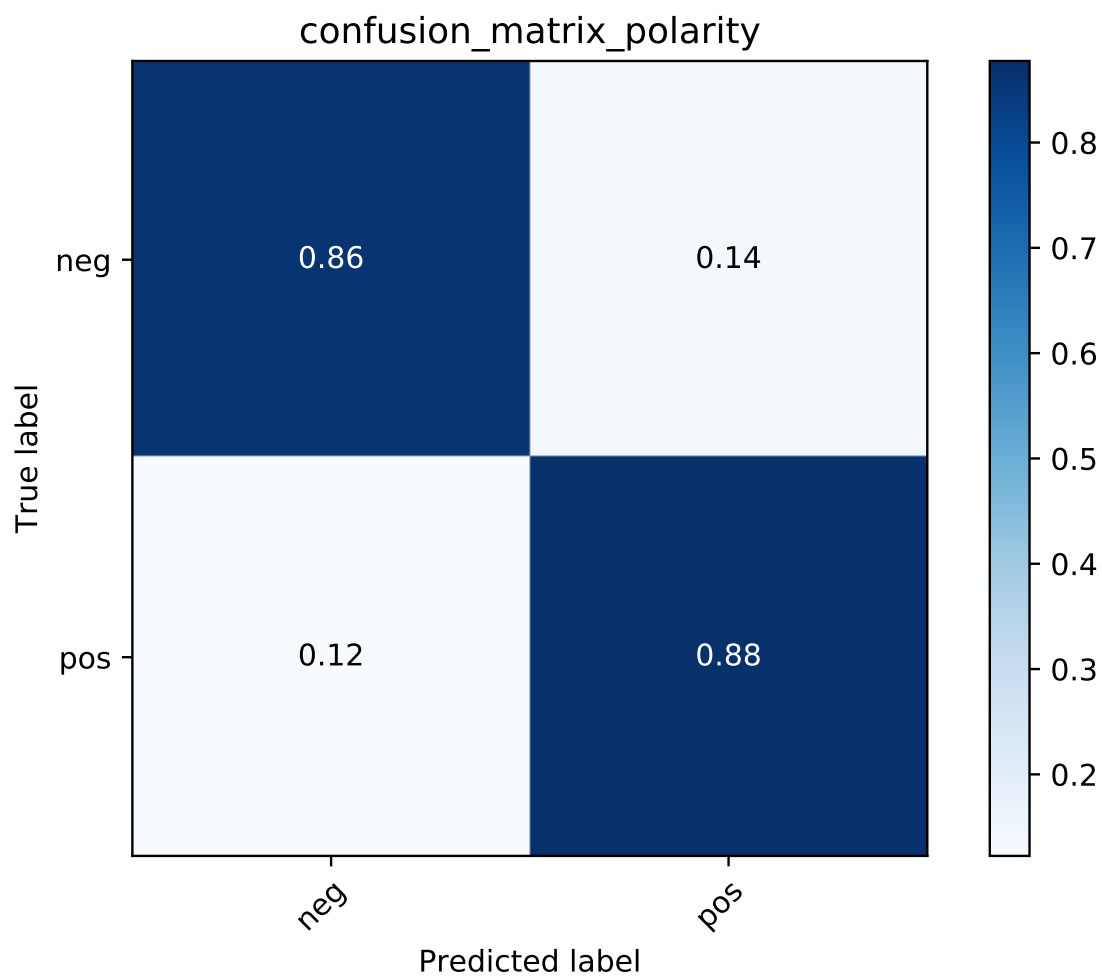


Figure 12: Polarity test confusion matrix for varying seeds.

Table 1: Sizes of seed word sets per class in fixed and varying regimes.

category	fixed seeds	varying seeds
+authority	40	54
+care	18	27
+fairness	18	26
+loyalty	16	30
+sanctity	21	33
-authority	18	32
-care	28	33
-fairness	12	23
-loyalty	7	16
-sanctity	30	53
pos	113	170
neg	95	157
moral	208	327
neutral	67	151

4 Prediction errors

Tables 2 and 3 show the mistakes made by the model in leave-one-out classification across all tasks in the decade of 1990, in the fixed seeds and varying seeds regimes, respectively.

Tables 4 and 5 show the same data for the decade of 1800.

Table 2: Prediction errors of centroid model on fixed seeds regime in 1990.

year	test	instance	true class	predicted class
1990	categorization	serve	+authority	+care
1990	categorization	respect	+authority	+fairness
1990	categorization	class	+authority	+loyalty
1990	categorization	leader	+authority	+loyalty
1990	categorization	mothers	+authority	+loyalty
1990	categorization	respected	+authority	+loyalty
1990	categorization	mother	+authority	+sanctity
1990	categorization	reverence	+authority	+sanctity
1990	categorization	tradition	+authority	+sanctity
1990	categorization	commit	+authority	-care
1990	categorization	lead	+authority	-care
1990	categorization	benefit	+care	+fairness
1990	categorization	compassion	+care	+sanctity
1990	categorization	compassionate	+care	+sanctity
1990	categorization	peace	+care	+sanctity
1990	categorization	guard	+care	-loyalty
1990	categorization	sympathy	+care	-sanctity

1990	categorization	rights	+fairness	+care
1990	categorization	justness	+fairness	-sanctity
1990	categorization	devote	+loyalty	+authority
1990	categorization	patriot	+loyalty	+sanctity
1990	categorization	unison	+loyalty	+sanctity
1990	categorization	clean	+sanctity	+care
1990	categorization	innocent	+sanctity	-care
1990	categorization	abstain	+sanctity	-fairness
1990	categorization	sanctify	+sanctity	-sanctity
1990	categorization	refuse	-authority	+authority
1990	categorization	alienate	-authority	+care
1990	categorization	obstruct	-authority	-care
1990	categorization	illegal	-authority	-fairness
1990	categorization	disobedient	-authority	-sanctity
1990	categorization	lawless	-authority	-sanctity
1990	categorization	sedition	-authority	-sanctity
1990	categorization	detriment	-care	+care
1990	categorization	impair	-care	+fairness
1990	categorization	annihilate	-care	-authority
1990	categorization	brutal	-care	-authority
1990	categorization	violence	-care	-authority
1990	categorization	wars	-care	-authority
1990	categorization	abuse	-care	-fairness
1990	categorization	kills	-care	-sanctity
1990	categorization	wound	-care	-sanctity
1990	categorization	preference	-fairness	+authority
1990	categorization	bias	-fairness	+fairness
1990	categorization	exclusive	-fairness	+fairness
1990	categorization	unjust	-fairness	-authority
1990	categorization	disproportion	-fairness	-sanctity
1990	categorization	injustice	-fairness	-sanctity
1990	categorization	foreign	-loyalty	+fairness
1990	categorization	individual	-loyalty	+loyalty
1990	categorization	spy	-loyalty	-authority
1990	categorization	terror	-loyalty	-authority
1990	categorization	deception	-loyalty	-care
1990	categorization	enemy	-loyalty	-care
1990	categorization	deceive	-loyalty	-sanctity
1990	categorization	sick	-sanctity	+care
1990	categorization	profane	-sanctity	+sanctity
1990	categorization	stain	-sanctity	+sanctity
1990	categorization	whore	-sanctity	+sanctity
1990	categorization	disease	-sanctity	-care
1990	categorization	repulse	-sanctity	-care
1990	categorization	gross	-sanctity	-fairness
1990	null test	amity	moral	neutral

1990	null'test	award	moral	neutral
1990	null'test	balance	moral	neutral
1990	null'test	bias	moral	neutral
1990	null'test	church	moral	neutral
1990	null'test	collective	moral	neutral
1990	null'test	command	moral	neutral
1990	null'test	companion	moral	neutral
1990	null'test	constant	moral	neutral
1990	null'test	dirt	moral	neutral
1990	null'test	disproportion	moral	neutral
1990	null'test	equal	moral	neutral
1990	null'test	equity	moral	neutral
1990	null'test	equivalent	moral	neutral
1990	null'test	exclusive	moral	neutral
1990	null'test	fairly	moral	neutral
1990	null'test	foreign	moral	neutral
1990	null'test	gross	moral	neutral
1990	null'test	group	moral	neutral
1990	null'test	hierarchy	moral	neutral
1990	null'test	joint	moral	neutral
1990	null'test	lead	moral	neutral
1990	null'test	member	moral	neutral
1990	null'test	position	moral	neutral
1990	null'test	preference	moral	neutral
1990	null'test	profligate	moral	neutral
1990	null'test	rank	moral	neutral
1990	null'test	reciprocal	moral	neutral
1990	null'test	refined	moral	neutral
1990	null'test	saint	moral	neutral
1990	null'test	serve	moral	neutral
1990	null'test	together	moral	neutral
1990	null'test	tradition	moral	neutral
1990	null'test	virgin	moral	neutral
1990	null'test	brazen	neutral	moral
1990	null'test	capricious	neutral	moral
1990	null'test	circumstance	neutral	moral
1990	null'test	customary	neutral	moral
1990	null'test	forge	neutral	moral
1990	null'test	induce	neutral	moral
1990	null'test	invisible	neutral	moral
1990	null'test	multitude	neutral	moral
1990	null'test	ought	neutral	moral
1990	null'test	pagan	neutral	moral
1990	null'test	piety	neutral	moral
1990	null'test	preposterous	neutral	moral
1990	null'test	priesthood	neutral	moral

1990	null'test	principle	neutral	moral
1990	null'test	repose	neutral	moral
1990	null'test	teeth	neutral	moral
1990	null'test	tempt	neutral	moral
1990	polarity	alienate	neg	pos
1990	polarity	defiance	neg	pos
1990	polarity	detriment	neg	pos
1990	polarity	disobedient	neg	pos
1990	polarity	dissent	neg	pos
1990	polarity	exclusion	neg	pos
1990	polarity	exclusive	neg	pos
1990	polarity	foreign	neg	pos
1990	polarity	individual	neg	pos
1990	polarity	preference	neg	pos
1990	polarity	prostitute	neg	pos
1990	polarity	refuse	neg	pos
1990	polarity	sick	neg	pos
1990	polarity	unequal	neg	pos
1990	polarity	whore	neg	pos
1990	polarity	abstain	pos	neg
1990	polarity	commit	pos	neg
1990	polarity	indifference	pos	neg
1990	polarity	innocent	pos	neg
1990	polarity	justifiable	pos	neg
1990	polarity	justifies	pos	neg
1990	polarity	justness	pos	neg
1990	polarity	lead	pos	neg
1990	polarity	sanctify	pos	neg
1990	polarity	shield	pos	neg
1990	polarity	sympathy	pos	neg
1990	polarity	virgins	pos	neg

Table 3: Prediction errors of centroid model on varying seeds regime in 1990.

year	test	instance	true' class	predicted' class
1990	categorization	respectful	+authority	+care
1990	categorization	serve	+authority	+care
1990	categorization	respect	+authority	+fairness
1990	categorization	bourgeoisie	+authority	+loyalty
1990	categorization	caste	+authority	+loyalty
1990	categorization	class	+authority	+loyalty
1990	categorization	leader	+authority	+loyalty
1990	categorization	leadership	+authority	+loyalty
1990	categorization	mothers	+authority	+loyalty

1990	categorization	respected	+authority	+loyalty
1990	categorization	tradition	+authority	+loyalty
1990	categorization	mother	+authority	+sanctity
1990	categorization	revere	+authority	+sanctity
1990	categorization	reverence	+authority	+sanctity
1990	categorization	commit	+authority	-care
1990	categorization	lead	+authority	-care
1990	categorization	mothering	+authority	-loyalty
1990	categorization	compliant	+authority	-sanctity
1990	categorization	deferent	+authority	-sanctity
1990	categorization	fatherly	+authority	-sanctity
1990	categorization	motherly	+authority	-sanctity
1990	categorization	obeisance	+authority	-sanctity
1990	categorization	venerate	+authority	-sanctity
1990	categorization	benefit	+care	+fairness
1990	categorization	amity	+care	+loyalty
1990	categorization	defence	+care	+loyalty
1990	categorization	defender	+care	+loyalty
1990	categorization	peace	+care	+loyalty
1990	categorization	sympathetic	+care	+loyalty
1990	categorization	compassionate	+care	+sanctity
1990	categorization	guard	+care	-care
1990	categorization	procurement	+care	-fairness
1990	categorization	pacify	+care	-sanctity
1990	categorization	sympathise	+care	-sanctity
1990	categorization	rights	+fairness	+care
1990	categorization	tolerant	+fairness	+care
1990	categorization	egalitarian	+fairness	+loyalty
1990	categorization	homologous	+fairness	+loyalty
1990	categorization	unprejudiced	+fairness	-authority
1990	categorization	justification	+fairness	-fairness
1990	categorization	justifies	+fairness	-fairness
1990	categorization	equable	+fairness	-sanctity
1990	categorization	evenness	+fairness	-sanctity
1990	categorization	justness	+fairness	-sanctity
1990	categorization	devote	+loyalty	+authority
1990	categorization	unison	+loyalty	+authority
1990	categorization	companionship	+loyalty	+care
1990	categorization	insider	+loyalty	+fairness
1990	categorization	patriot	+loyalty	-authority
1990	categorization	unite	+loyalty	-authority
1990	categorization	clique	+loyalty	-fairness
1990	categorization	communism	+loyalty	-loyalty
1990	categorization	comrade	+loyalty	-sanctity
1990	categorization	clean	+sanctity	+care
1990	categorization	innocent	+sanctity	-care

1990	categorization	abstain	+sanctity	-fairness
1990	categorization	abstention	+sanctity	-fairness
1990	categorization	abstinent	+sanctity	-sanctity
1990	categorization	limpid	+sanctity	-sanctity
1990	categorization	refined	+sanctity	-sanctity
1990	categorization	saintly	+sanctity	-sanctity
1990	categorization	sanctify	+sanctity	-sanctity
1990	categorization	sterile	+sanctity	-sanctity
1990	categorization	unadulterated	+sanctity	-sanctity
1990	categorization	disobey	-authority	+authority
1990	categorization	refuse	-authority	+authority
1990	categorization	alienate	-authority	+care
1990	categorization	dissident	-authority	+loyalty
1990	categorization	obstruct	-authority	-care
1990	categorization	riot	-authority	-care
1990	categorization	illegal	-authority	-fairness
1990	categorization	illegality	-authority	-fairness
1990	categorization	agitator	-authority	-loyalty
1990	categorization	insurgent	-authority	-loyalty
1990	categorization	disobedient	-authority	-sanctity
1990	categorization	lawless	-authority	-sanctity
1990	categorization	mutinous	-authority	-sanctity
1990	categorization	nonconformist	-authority	-sanctity
1990	categorization	remonstrate	-authority	-sanctity
1990	categorization	riotous	-authority	-sanctity
1990	categorization	unfaithful	-authority	-sanctity
1990	categorization	detriment	-care	+care
1990	categorization	detrimental	-care	+care
1990	categorization	impair	-care	+fairness
1990	categorization	warring	-care	+loyalty
1990	categorization	annihilate	-care	-authority
1990	categorization	brutal	-care	-authority
1990	categorization	wars	-care	-loyalty
1990	categorization	kills	-care	-sanctity
1990	categorization	ravage	-care	-sanctity
1990	categorization	spurn	-care	-sanctity
1990	categorization	warlord	-care	-sanctity
1990	categorization	preference	-fairness	+authority
1990	categorization	exclusive	-fairness	+fairness
1990	categorization	dishonest	-fairness	-authority
1990	categorization	unjust	-fairness	-authority
1990	categorization	unscrupulous	-fairness	-authority
1990	categorization	bigot	-fairness	-sanctity
1990	categorization	discriminator	-fairness	-sanctity
1990	categorization	disproportion	-fairness	-sanctity
1990	categorization	dissociate	-fairness	-sanctity

1990	categorization	favoritism	-fairness	-sanctity
1990	categorization	prejudge	-fairness	-sanctity
1990	categorization	foreign	-loyalty	+loyalty
1990	categorization	immigrant	-loyalty	+loyalty
1990	categorization	individual	-loyalty	+loyalty
1990	categorization	immigrate	-loyalty	-authority
1990	categorization	spy	-loyalty	-authority
1990	categorization	terror	-loyalty	-authority
1990	categorization	deception	-loyalty	-care
1990	categorization	enemy	-loyalty	-care
1990	categorization	terrorist	-loyalty	-care
1990	categorization	deceptive	-loyalty	-fairness
1990	categorization	personhood	-loyalty	-fairness
1990	categorization	deceive	-loyalty	-sanctity
1990	categorization	miscreant	-loyalty	-sanctity
1990	categorization	personify	-loyalty	-sanctity
1990	categorization	renegade	-loyalty	-sanctity
1990	categorization	exploit	-sanctity	+authority
1990	categorization	sick	-sanctity	+care
1990	categorization	lax	-sanctity	+fairness
1990	categorization	profane	-sanctity	+sanctity
1990	categorization	prostitute	-sanctity	+sanctity
1990	categorization	stain	-sanctity	+sanctity
1990	categorization	whore	-sanctity	+sanctity
1990	categorization	desecration	-sanctity	-authority
1990	categorization	obscenity	-sanctity	-authority
1990	categorization	disease	-sanctity	-care
1990	categorization	repulse	-sanctity	-care
1990	categorization	exploitation	-sanctity	-fairness
1990	categorization	gross	-sanctity	-fairness
1990	categorization	obscene	-sanctity	-fairness
1990	null test	abstinent	moral	neutral
1990	null test	agitate	moral	neutral
1990	null test	agitator	moral	neutral
1990	null test	award	moral	neutral
1990	null test	balance	moral	neutral
1990	null test	battle	moral	neutral
1990	null test	bigot	moral	neutral
1990	null test	cadre	moral	neutral
1990	null test	clean	moral	neutral
1990	null test	clique	moral	neutral
1990	null test	cohort	moral	neutral
1990	null test	command	moral	neutral
1990	null test	companion	moral	neutral
1990	null test	compliant	moral	neutral
1990	null test	constant	moral	neutral

1990	null test	contaminate	moral	neutral
1990	null test	corruptible	moral	neutral
1990	null test	deferent	moral	neutral
1990	null test	dirt	moral	neutral
1990	null test	discriminative	moral	neutral
1990	null test	discriminator	moral	neutral
1990	null test	disproportion	moral	neutral
1990	null test	equable	moral	neutral
1990	null test	equivalent	moral	neutral
1990	null test	evenness	moral	neutral
1990	null test	fairly	moral	neutral
1990	null test	fatherly	moral	neutral
1990	null test	filth	moral	neutral
1990	null test	gross	moral	neutral
1990	null test	group	moral	neutral
1990	null test	guild	moral	neutral
1990	null test	hierarchical	moral	neutral
1990	null test	homologous	moral	neutral
1990	null test	insurgent	moral	neutral
1990	null test	joint	moral	neutral
1990	null test	lead	moral	neutral
1990	null test	limpid	moral	neutral
1990	null test	maiden	moral	neutral
1990	null test	motherly	moral	neutral
1990	null test	mutinous	moral	neutral
1990	null test	obeisance	moral	neutral
1990	null test	permission	moral	neutral
1990	null test	personify	moral	neutral
1990	null test	prejudge	moral	neutral
1990	null test	pristine	moral	neutral
1990	null test	procurement	moral	neutral
1990	null test	profligate	moral	neutral
1990	null test	ravage	moral	neutral
1990	null test	rebuff	moral	neutral
1990	null test	reciprocal	moral	neutral
1990	null test	refined	moral	neutral
1990	null test	repulse	moral	neutral
1990	null test	revere	moral	neutral
1990	null test	riotous	moral	neutral
1990	null test	saint	moral	neutral
1990	null test	saintly	moral	neutral
1990	null test	shield	moral	neutral
1990	null test	slut	moral	neutral
1990	null test	spurn	moral	neutral
1990	null test	spy	moral	neutral
1990	null test	stain	moral	neutral

1990	null test	sterile	moral	neutral
1990	null test	sterility	moral	neutral
1990	null test	sympathise	moral	neutral
1990	null test	tarnish	moral	neutral
1990	null test	together	moral	neutral
1990	null test	tramp	moral	neutral
1990	null test	unadulterated	moral	neutral
1990	null test	unbiased	moral	neutral
1990	null test	venerate	moral	neutral
1990	null test	virgin	moral	neutral
1990	null test	warlord	moral	neutral
1990	null test	wound	moral	neutral
1990	null test	capricious	neutral	moral
1990	null test	circumstance	neutral	moral
1990	null test	circumvent	neutral	moral
1990	null test	contender	neutral	moral
1990	null test	context	neutral	moral
1990	null test	convene	neutral	moral
1990	null test	convention	neutral	moral
1990	null test	customary	neutral	moral
1990	null test	depend	neutral	moral
1990	null test	diagnose	neutral	moral
1990	null test	figment	neutral	moral
1990	null test	financial	neutral	moral
1990	null test	forge	neutral	moral
1990	null test	franchise	neutral	moral
1990	null test	indication	neutral	moral
1990	null test	induce	neutral	moral
1990	null test	insistence	neutral	moral
1990	null test	interstate	neutral	moral
1990	null test	invisible	neutral	moral
1990	null test	minister	neutral	moral
1990	null test	multitude	neutral	moral
1990	null test	ought	neutral	moral
1990	null test	pact	neutral	moral
1990	null test	pagan	neutral	moral
1990	null test	piety	neutral	moral
1990	null test	preposterous	neutral	moral
1990	null test	priesthood	neutral	moral
1990	null test	principle	neutral	moral
1990	null test	psyche	neutral	moral
1990	null test	relentless	neutral	moral
1990	null test	repose	neutral	moral
1990	null test	ruler	neutral	moral
1990	null test	tempt	neutral	moral
1990	null test	testament	neutral	moral

1990	polarity	alienate	neg	pos
1990	polarity	bias	neg	pos
1990	polarity	corruptible	neg	pos
1990	polarity	defiance	neg	pos
1990	polarity	detriment	neg	pos
1990	polarity	detrimental	neg	pos
1990	polarity	discriminative	neg	pos
1990	polarity	disobedient	neg	pos
1990	polarity	dissent	neg	pos
1990	polarity	exclusion	neg	pos
1990	polarity	exclusive	neg	pos
1990	polarity	exploit	neg	pos
1990	polarity	foreign	neg	pos
1990	polarity	gross	neg	pos
1990	polarity	immigrant	neg	pos
1990	polarity	immigrate	neg	pos
1990	polarity	individual	neg	pos
1990	polarity	insurgent	neg	pos
1990	polarity	lax	neg	pos
1990	polarity	personhood	neg	pos
1990	polarity	preference	neg	pos
1990	polarity	prostitute	neg	pos
1990	polarity	refuse	neg	pos
1990	polarity	sick	neg	pos
1990	polarity	unequal	neg	pos
1990	polarity	unfaithful	neg	pos
1990	polarity	warring	neg	pos
1990	polarity	whore	neg	pos
1990	polarity	abstain	pos	neg
1990	polarity	abstention	pos	neg
1990	polarity	abstinent	pos	neg
1990	polarity	bourgeoisie	pos	neg
1990	polarity	chaste	pos	neg
1990	polarity	clique	pos	neg
1990	polarity	commit	pos	neg
1990	polarity	communism	pos	neg
1990	polarity	compliant	pos	neg
1990	polarity	comrade	pos	neg
1990	polarity	deferent	pos	neg
1990	polarity	equable	pos	neg
1990	polarity	evenness	pos	neg
1990	polarity	fatherly	pos	neg
1990	polarity	indifference	pos	neg
1990	polarity	innocent	pos	neg
1990	polarity	insider	pos	neg
1990	polarity	justifiable	pos	neg

1990	polarity	justifies	pos	neg
1990	polarity	justness	pos	neg
1990	polarity	limpid	pos	neg
1990	polarity	mothering	pos	neg
1990	polarity	obeisance	pos	neg
1990	polarity	pacify	pos	neg
1990	polarity	revere	pos	neg
1990	polarity	saintly	pos	neg
1990	polarity	sanctify	pos	neg
1990	polarity	shield	pos	neg
1990	polarity	sterile	pos	neg
1990	polarity	sympathise	pos	neg
1990	polarity	unadulterated	pos	neg
1990	polarity	unprejudiced	pos	neg
1990	polarity	venerate	pos	neg
1990	polarity	virgins	pos	neg

Table 4: Prediction errors of centroid model on fixed seeds regime in 1800.

year	test	instance	true class	predicted class
1800	categorization	award	+authority	+fairness
1800	categorization	position	+authority	+fairness
1800	categorization	class	+authority	+loyalty
1800	categorization	leader	+authority	+loyalty
1800	categorization	mothers	+authority	+sanctity
1800	categorization	tradition	+authority	+sanctity
1800	categorization	hierarchy	+authority	-authority
1800	categorization	commit	+authority	-sanctity
1800	categorization	lead	+authority	-sanctity
1800	categorization	amity	+care	+loyalty
1800	categorization	benefit	+care	+loyalty
1800	categorization	protector	+care	+loyalty
1800	categorization	sympathy	+care	+loyalty
1800	categorization	compassion	+care	+sanctity
1800	categorization	defender	+care	+sanctity
1800	categorization	guard	+care	-loyalty
1800	categorization	compassionate	+care	-sanctity
1800	categorization	constant	+fairness	+authority
1800	categorization	rights	+fairness	+care
1800	categorization	justifiable	+fairness	-fairness
1800	categorization	reciprocal	+fairness	-fairness
1800	categorization	devote	+loyalty	+authority
1800	categorization	ally	+loyalty	+care
1800	categorization	companion	+loyalty	+sanctity

1800	categorization	unison	+loyalty	+sanctity
1800	categorization	church	+sanctity	+authority
1800	categorization	sacred	+sanctity	+authority
1800	categorization	virtuous	+sanctity	+loyalty
1800	categorization	abstain	+sanctity	-authority
1800	categorization	clean	+sanctity	-sanctity
1800	categorization	innocent	+sanctity	-sanctity
1800	categorization	refined	+sanctity	-sanctity
1800	categorization	sanctify	+sanctity	-sanctity
1800	categorization	refuse	-authority	+authority
1800	categorization	alienate	-authority	-care
1800	categorization	obstruct	-authority	-care
1800	categorization	disobedience	-authority	-sanctity
1800	categorization	disobedient	-authority	-sanctity
1800	categorization	riot	-authority	-sanctity
1800	categorization	annihilate	-care	-authority
1800	categorization	crush	-care	-authority
1800	categorization	detrimental	-care	-fairness
1800	categorization	impair	-care	-fairness
1800	categorization	abuse	-care	-sanctity
1800	categorization	brutal	-care	-sanctity
1800	categorization	cruel	-care	-sanctity
1800	categorization	kills	-care	-sanctity
1800	categorization	bias	-fairness	+authority
1800	categorization	exclusive	-fairness	+fairness
1800	categorization	unequal	-fairness	+fairness
1800	categorization	unfair	-fairness	+fairness
1800	categorization	injustice	-fairness	-sanctity
1800	categorization	foreign	-loyalty	+care
1800	categorization	individual	-loyalty	+loyalty
1800	categorization	enemy	-loyalty	-care
1800	categorization	terror	-loyalty	-care
1800	categorization	deceive	-loyalty	-sanctity
1800	categorization	deception	-loyalty	-sanctity
1800	categorization	spy	-loyalty	-sanctity
1800	categorization	blemish	-sanctity	+sanctity
1800	categorization	impious	-sanctity	-authority
1800	categorization	intemperate	-sanctity	-authority
1800	categorization	repulse	-sanctity	-authority
1800	categorization	disease	-sanctity	-care
1800	categorization	sick	-sanctity	-care
1800	categorization	disgust	-sanctity	-fairness
1800	categorization	pervert	-sanctity	-fairness
1800	null test	balance	moral	neutral
1800	null test	battle	moral	neutral
1800	null test	class	moral	neutral

1800	null'test	clean	moral	neutral
1800	null'test	deception	moral	neutral
1800	null'test	defile	moral	neutral
1800	null'test	dirt	moral	neutral
1800	null'test	discrimination	moral	neutral
1800	null'test	disease	moral	neutral
1800	null'test	disproportion	moral	neutral
1800	null'test	equal	moral	neutral
1800	null'test	equivalent	moral	neutral
1800	null'test	exclusive	moral	neutral
1800	null'test	fair	moral	neutral
1800	null'test	fairly	moral	neutral
1800	null'test	filth	moral	neutral
1800	null'test	gross	moral	neutral
1800	null'test	group	moral	neutral
1800	null'test	joint	moral	neutral
1800	null'test	justness	moral	neutral
1800	null'test	lead	moral	neutral
1800	null'test	member	moral	neutral
1800	null'test	permission	moral	neutral
1800	null'test	position	moral	neutral
1800	null'test	pure	moral	neutral
1800	null'test	refined	moral	neutral
1800	null'test	saint	moral	neutral
1800	null'test	shield	moral	neutral
1800	null'test	together	moral	neutral
1800	null'test	tradition	moral	neutral
1800	null'test	antiquated	neutral	moral
1800	null'test	capricious	neutral	moral
1800	null'test	circumstance	neutral	moral
1800	null'test	council	neutral	moral
1800	null'test	credit	neutral	moral
1800	null'test	customary	neutral	moral
1800	null'test	depend	neutral	moral
1800	null'test	franchise	neutral	moral
1800	null'test	indication	neutral	moral
1800	null'test	induce	neutral	moral
1800	null'test	invisible	neutral	moral
1800	null'test	minister	neutral	moral
1800	null'test	multitude	neutral	moral
1800	null'test	ought	neutral	moral
1800	null'test	pagan	neutral	moral
1800	null'test	piety	neutral	moral
1800	null'test	preposterous	neutral	moral
1800	null'test	priesthood	neutral	moral
1800	null'test	principle	neutral	moral

1800	null'test	repose	neutral	moral
1800	null'test	ruler	neutral	moral
1800	null'test	tempt	neutral	moral
1800	polarity	bias	neg	pos
1800	polarity	defiance	neg	pos
1800	polarity	detrimental	neg	pos
1800	polarity	discrimination	neg	pos
1800	polarity	dissent	neg	pos
1800	polarity	exclusion	neg	pos
1800	polarity	exclusive	neg	pos
1800	polarity	foreign	neg	pos
1800	polarity	individual	neg	pos
1800	polarity	preference	neg	pos
1800	polarity	protest	neg	pos
1800	polarity	refuse	neg	pos
1800	polarity	sinner	neg	pos
1800	polarity	sins	neg	pos
1800	polarity	unequal	neg	pos
1800	polarity	abstain	pos	neg
1800	polarity	commit	pos	neg
1800	polarity	fairly	pos	neg
1800	polarity	guard	pos	neg
1800	polarity	innocent	pos	neg
1800	polarity	justifiable	pos	neg
1800	polarity	justifies	pos	neg
1800	polarity	shelter	pos	neg

Table 5: Prediction errors of centroid model on varying seeds regime in 1800.

year	test	instance	true' class	predicted' class
1800	categorization	award	+authority	+fairness
1800	categorization	position	+authority	+fairness
1800	categorization	class	+authority	+loyalty
1800	categorization	leader	+authority	+loyalty
1800	categorization	mothers	+authority	+sanctity
1800	categorization	tradition	+authority	+sanctity
1800	categorization	hierarchy	+authority	-authority
1800	categorization	commit	+authority	-sanctity
1800	categorization	lead	+authority	-sanctity
1800	categorization	amity	+care	+loyalty
1800	categorization	benefit	+care	+loyalty
1800	categorization	protector	+care	+loyalty
1800	categorization	sympathy	+care	+loyalty
1800	categorization	compassion	+care	+sanctity

1800	categorization	defender	+care	+sanctity
1800	categorization	guard	+care	-loyalty
1800	categorization	compassionate	+care	-sanctity
1800	categorization	constant	+fairness	+authority
1800	categorization	rights	+fairness	+care
1800	categorization	justifiable	+fairness	-fairness
1800	categorization	reciprocal	+fairness	-fairness
1800	categorization	devote	+loyalty	+authority
1800	categorization	ally	+loyalty	+care
1800	categorization	companion	+loyalty	+sanctity
1800	categorization	unison	+loyalty	+sanctity
1800	categorization	church	+sanctity	+authority
1800	categorization	sacred	+sanctity	+authority
1800	categorization	virtuous	+sanctity	+loyalty
1800	categorization	abstain	+sanctity	-authority
1800	categorization	clean	+sanctity	-sanctity
1800	categorization	innocent	+sanctity	-sanctity
1800	categorization	refined	+sanctity	-sanctity
1800	categorization	sanctify	+sanctity	-sanctity
1800	categorization	refuse	-authority	+authority
1800	categorization	alienate	-authority	-care
1800	categorization	obstruct	-authority	-care
1800	categorization	disobedience	-authority	-sanctity
1800	categorization	disobedient	-authority	-sanctity
1800	categorization	riot	-authority	-sanctity
1800	categorization	annihilate	-care	-authority
1800	categorization	crush	-care	-authority
1800	categorization	detrimental	-care	-fairness
1800	categorization	impair	-care	-fairness
1800	categorization	abuse	-care	-sanctity
1800	categorization	brutal	-care	-sanctity
1800	categorization	cruel	-care	-sanctity
1800	categorization	kills	-care	-sanctity
1800	categorization	bias	-fairness	+authority
1800	categorization	exclusive	-fairness	+fairness
1800	categorization	unequal	-fairness	+fairness
1800	categorization	unfair	-fairness	+fairness
1800	categorization	injustice	-fairness	-sanctity
1800	categorization	foreign	-loyalty	+care
1800	categorization	individual	-loyalty	+loyalty
1800	categorization	enemy	-loyalty	-care
1800	categorization	terror	-loyalty	-care
1800	categorization	deceive	-loyalty	-sanctity
1800	categorization	deception	-loyalty	-sanctity
1800	categorization	spy	-loyalty	-sanctity
1800	categorization	blemish	-sanctity	+sanctity

1800	categorization	impious	-sanctity	-authority
1800	categorization	intemperate	-sanctity	-authority
1800	categorization	repulse	-sanctity	-authority
1800	categorization	disease	-sanctity	-care
1800	categorization	sick	-sanctity	-care
1800	categorization	disgust	-sanctity	-fairness
1800	categorization	pervert	-sanctity	-fairness
1800	null'test	balance	moral	neutral
1800	null'test	battle	moral	neutral
1800	null'test	class	moral	neutral
1800	null'test	clean	moral	neutral
1800	null'test	deception	moral	neutral
1800	null'test	defile	moral	neutral
1800	null'test	dirt	moral	neutral
1800	null'test	discrimination	moral	neutral
1800	null'test	disease	moral	neutral
1800	null'test	disproportion	moral	neutral
1800	null'test	equal	moral	neutral
1800	null'test	equivalent	moral	neutral
1800	null'test	exclusive	moral	neutral
1800	null'test	fair	moral	neutral
1800	null'test	fairly	moral	neutral
1800	null'test	filth	moral	neutral
1800	null'test	gross	moral	neutral
1800	null'test	group	moral	neutral
1800	null'test	joint	moral	neutral
1800	null'test	justness	moral	neutral
1800	null'test	lead	moral	neutral
1800	null'test	member	moral	neutral
1800	null'test	permission	moral	neutral
1800	null'test	position	moral	neutral
1800	null'test	pure	moral	neutral
1800	null'test	refined	moral	neutral
1800	null'test	saint	moral	neutral
1800	null'test	shield	moral	neutral
1800	null'test	together	moral	neutral
1800	null'test	tradition	moral	neutral
1800	null'test	antiquated	neutral	moral
1800	null'test	borough	neutral	moral
1800	null'test	capricious	neutral	moral
1800	null'test	circumstance	neutral	moral
1800	null'test	council	neutral	moral
1800	null'test	credit	neutral	moral
1800	null'test	customary	neutral	moral
1800	null'test	depend	neutral	moral
1800	null'test	franchise	neutral	moral

1800	null'test	indication	neutral	moral
1800	null'test	induce	neutral	moral
1800	null'test	invisible	neutral	moral
1800	null'test	minister	neutral	moral
1800	null'test	multitude	neutral	moral
1800	null'test	ought	neutral	moral
1800	null'test	pagan	neutral	moral
1800	null'test	piety	neutral	moral
1800	null'test	preposterous	neutral	moral
1800	null'test	priesthood	neutral	moral
1800	null'test	principle	neutral	moral
1800	null'test	repose	neutral	moral
1800	null'test	ruler	neutral	moral
1800	null'test	tempt	neutral	moral
1800	polarity	bias	neg	pos
1800	polarity	defiance	neg	pos
1800	polarity	detrimental	neg	pos
1800	polarity	discrimination	neg	pos
1800	polarity	dissent	neg	pos
1800	polarity	exclusion	neg	pos
1800	polarity	exclusive	neg	pos
1800	polarity	foreign	neg	pos
1800	polarity	individual	neg	pos
1800	polarity	preference	neg	pos
1800	polarity	protest	neg	pos
1800	polarity	refuse	neg	pos
1800	polarity	sinner	neg	pos
1800	polarity	sins	neg	pos
1800	polarity	unequal	neg	pos
1800	polarity	abstain	pos	neg
1800	polarity	commit	pos	neg
1800	polarity	fairly	pos	neg
1800	polarity	guard	pos	neg
1800	polarity	innocent	pos	neg
1800	polarity	justifiable	pos	neg
1800	polarity	justifies	pos	neg
1800	polarity	shelter	pos	neg