Disgust is a factor in extreme prejudice

Kathleen Taylor*
University of Oxford, UK

Understanding intergroup prejudice is a dominant research focus for social psychology. Prejudice is usually conceptualized as a continuum of positive/negative affect, but this has limitations. It neither reflects people's ability to maintain simultaneous positive and negative stereotypes of others nor explains extreme prejudice (bigotry). Some researchers have proposed multidimensional models of prejudice in which different negative emotions are evoked depending on the situation. Extending this to bigotry raises the question of which emotions are most relevant. Therefore, this study looked at 'anti-group' texts – writings which promote extreme intergroup hostility – and analysed the frequency of emotive language. Findings suggest that bigotry may be distinguished by high levels of disgust.

Within social psychology, the study of intergroup relations has been dominated by the concept of prejudice, and by a cognitive, attitude-based theoretical framework which views prejudice as the favourable or unfavourable evaluation of a group and its members. This framework, which has been highly successful, conceptualizes prejudice as a unidimensional evaluative continuum (Allport, 1954; Billig, 2002a; Brewer & Brown, 1998; Tajfel, 1969). Through learning and socialization processes, positive or negative affect becomes attached to one person or group's mental representation of one another.

More recently, however, some limitations of the cognitive approach have begun to emerge. One is that conceptualizing prejudice as an attitude implies an affective stability across circumstances which does not seem to reflect observable variability in how groups are evaluated. Accordingly, some researchers have begun to differentiate emotions within prejudice and explore the hypothesis that prejudice against different groups may have distinct emotional profiles (for recent reviews see Cottrell & Neuberg, 2005; Mackie & Smith, 2002). Fiske and colleagues (Fiske, Cuddy, & Glick, 2002; Fiske, Cuddy, Glick, & Xu, 2002; Fiske, Xu, Cuddy, & Glick, 1999) have shown that American stereotypes of some minority groups (e.g. Asian-Americans, Jewish people, career women) involve both positive emotions (such as respect/admiration) and negative emotions (such as dislike/ill will). The positive evaluations seem to be linked to these groups’ perceived high status, while the negative evaluations appear to result from their
being viewed as potential or actual competitors. Making issues of status or competition
more or less salient can therefore alter the degree of affect attached to a group’s
stereotype, implying more flexibility than is traditionally associated with attitudes.

A second limitation of the cognitive approach is that it does not differentiate within
affect, other than to discriminate between negative and positive affects. It therefore
lumps together negative emotions, such as anger, fear, disgust and sadness. Yet these are
qualitatively and physiologically different (Collet, Vernet-Maury, Dellhomme, & Dittmar,
1997; Darwin, 1999; Ekman, 1999; Esslen, Pascual-Marqui, Hell, Kochi, & Lehmann,
2004), occur under different circumstances and facilitate different and sometimes
contradictory responses (Cottrell & Neuberg, 2005; Mackie & Smith, 2002). Anger
‘produces a tendency to perceive negative events as predictable, under human control
and brought about by others, while fear produces a tendency to perceive negative
events as unpredictable and under situational control’ (Fessler, Pillsworth, & Flannery,
2004, p. 108). Anger thus tends to prompt aggression, while fear prompts
avoidance/escape; sadness prompts withdrawal. Disgust may prompt either avoidance
or attempts to remove the disgusting stimulus, depending on which is easier to achieve
(Cottrell & Neuberg, 2005).

A third criticism of the cognitive approach to prejudice is that it fails to explain the
extremes of bigotry sometimes observed in the world but rarely in the laboratory (Billig,
2002a, 2002b; Brown, 2002; Frosh, 2002). As Dutton and colleagues conclude with
reference to soldiers who massacre, ‘the explanation of the specific forms of violence,
rape, mutilation, torture, etc., is not forthcoming from current psychological knowledge
[. . .] psychology has not attempted to account for the extremity of massacre’ (Dutton,
Boyanowsky, & Bond, 2005, p. 470). Conceptualizing prejudice as a continuum may be
taken as implying that progression towards the extremes of that continuum simply
involves ‘more of the same’ contributory factors, rather than a qualitative shift in which
new factors come into play. But this is contentious: mild dislike seems very different
from the murderous hatred often expressed during atrocities, for example, by the radio
station Radio Télévision Libre des Milles Collines (RTLM) during the Rwandan genocide
(Dallaire, 2004).

The language of bigotry
Understanding prejudice and bigotry seems to require moving beyond a unidimensional
conception of affect to differentiate negative emotions. One approach is to look at
language; specifically, the written language of political extremists: ‘anti-group’ texts
such as Adolf Hitler’s Mein Kampf (Hitler, trans. 1939) in which the authors target,
denigrate and express extreme hostility to a particular out-group or out-groups and their
members. If there is a particular emotion profile (pattern of usage of emotion-related
words) which characterizes anti-group texts and distinguishes them from language in
general, then this may be quantifiable, in the form of differences in the prevalences of
words associated with specific emotions (hereafter ‘E-words’).

Two assumptions underlie this hypothesis. The first is that the frequency with which
an E-word is used in the text reflects its relevance to the text’s social function as
perceived by the author(s). ‘People use their language to do things’ (Potter & Wetherell,
1987, p. 32, original italics). Anti-group texts provide information; however, they also
attempt to manipulate their readers’ beliefs by asserting, implying and/or justifying
negative claims about people, processes which require the association of cognitions
about the target group with negatively valenced emotive words. The more frequently
such links are made, the greater is their importance to the arguments put forward in the
text. An example from one of the texts used in this study is the association of Islam and
Muslims with terrorism (Barnes, 2004). In a text of 1,664 words, the words ‘terror’,
‘terrorism’ and ‘terrorist(s)’ appear 35 times and words describing the target group
(‘Islam/ic’, ‘mosque’, ‘Muslim’) occur 37 times, more frequently than any other words
except for common filler words like ‘and’ (50 occurrences). The two are linked directly
(within a single noun phrase) on five occasions. The text never explicitly makes the
generic claim that ‘Muslims are terrorists’; it does not need to.

The second assumption is that the authors of anti-group texts use E-words with the
aim of evoking the corresponding emotions in their presumed readers, who for anti-
group texts tend to be members of the authors’ in-group (i.e. ideological sympathizers).
The presence of E-words in texts may not bear a simple relation to the authors’ own
feelings – for example because of authorial lack of insight (Blascovich, Mendes, & Seery,
2002) – nevertheless, the assumption is that the function of a phrase such as ‘crawled
like cockroaches’ (Barnes, 2004) is to arouse disgust, however mild and fleeting, in the
reader.

Does the strategy work? Inferring audience effects from anti-group texts is
problematic for several reasons. First, one needs to specify the effects in question,
which may be emotional, behavioural or facilitatory. In principle, emotional effects may
range from the negligible - or even the counterproductive arousal of resistance in
readers who resent being manipulated - to significant emotional responses; while
behavioural effects may range from the minor (e.g. adopting phrases from the text) to
the life-changing. Facilitatory effects do not directly induce hostile behaviour, but they
enhance motives to perform it and reduce the inhibitions which generally prevent it
(Taylor, 2006), for example by legitimating the expression of negative emotions directed
at the target group (Yzerbyt, Schadron, Leyens, & Rocher, 1994). In practice, facilitatory
effects are more likely than behavioural responses (Taylor, 2006).

More controversial is the claim that the presence of E-words in the text may evoke
emotional responses in the reader, especially given that the method used in this study
assesses the frequencies of words in isolation from their semantic context. As discourse
analysts have noted, word meaning may vary with context, both textual and social
(Potter & Wetherell, 1987). However, three points should be noted in relation to this
study. First, variations in the meaning of an E-word are less important than variations in
the type of valence. Thus, semantic contexts which modify valence intensity, for
example by adding a qualifier such as ‘mild’ or ‘extreme’, can be regarded as valence-
congruent. Incongruent contexts are those which change the sign of the word’s
valence, for example by negation. The semantic context of every E-word occurrence in
the texts used in this study was therefore assessed for congruence (see Results section).

Secondly, recent evidence supports the claim that E-words can evoke emotional
responses. Neuroimaging studies indicate that reading descriptions of disgusting
scenarios, perceiving disgusted facial expressions and viewing disgusting images
activate similar brain areas (Moll et al., 2005; Phan, Wager, Taylor, & Liberzon, 2002).
Emotional responses to stimuli may be very fast (Esslen et al., 2004), suggesting that
they may occur prior to the comprehension of semantic context. In addition, a recent
study by Wheatley and Haidt (2005) used hypnosis to associate semantically
unconnected words (‘often’ or ‘take’) with feelings of disgust. Post-hypnosis, the
participants were asked to rate the severity and disgustingness of moral transgressions
described in vignettes which were identical except that they included either the word
‘often’ or the word ‘take’. The mere presence of a word hypnotically linked to disgust
significantly increased disgust ratings and the severity of moral judgments, even when participants recognized the unreasonableness of those judgments. Similarly, Lerner, Small, and Loewenstein (2004) demonstrate that disgust aroused by an irrelevant situation ‘can have dramatic effects on economic transactions’ (p. 339).

Thirdly, in this study, the emotional impact of the anti-group texts was assessed directly by presenting them to two independent raters, who were asked how emotive they found the texts. They were also asked to generate, for each text, a list of any emotion which they felt the text expressed or which the author was trying to convey (see Results section).

Emotions in atrocities
If anti-group texts show differentiated emotion profiles relative to language in general, which negative emotions should be relevant? Historians, cultural theorists and psychologists who have studied genocides, mass killings and other atrocities have often noted the emotive language used by perpetrators and their sympathizers to describe members of the victimized out-group and advocate action against them. Yet research has tended to focus on cognitive phenomena, such as the relative importance of ideological propaganda and the already-present cultural beliefs (Goldhagen, 1997; Koch, 2000) or the roles of stereotyping, obedience to authority and conformity (Browning, 1991; Milgram, 1997; Newman & Erber, 2002). Less attention has been paid to the question of which emotions are emphasized by perpetrators themselves, for example in their communications with their followers. The consensus with respect to genocide, insofar as there is one, seems to be that (at least) anger and frustration, disgust, envy, fear and hate may all play a role (Fein, 1990; Lemarchand, 2002; Mandel, 2002; Newman & Erber, 2002).

Social psychological research on prejudice makes varying predictions of which negative emotions should predominate. Integrated threat theory, for example, frames prejudice as a reaction to threat and thus views fear and anger as the dominant, although not the only, emotions involved (Stephan & Renfro, 2002; Stephan & Stephan, 2000). Since bigotry is more closely linked to aggression than is mild prejudice, anger rather than fear may dominate the emotion profile (Cottrell & Neuberg, 2005).

Fiske and colleagues (1999; Harris & Fiske, 2006) link emotion profiles in prejudice to the targeted group’s perceived status and competence. High-status groups evoke either pride and admiration (if they are liked, i.e. not seen as potential competitors) or envy, jealousy, hostility and depression (if they are competitors), while low-status groups evoke pity and sympathy (if they are liked) or contempt, resentment, hate, anger and disgust (if they are disliked). Most victims of atrocities are from groups seen as low status by the aggressor because of pervasive cultural attitudes – for example women and children, the elderly, the physically or mentally ill or disabled – or specific beliefs – for example Jews under the Nazi regime, Armenian Christians in Turkey during the First World War, Bosnian Muslims in the disintegration of Yugoslavia and ethnic Chinese during the capture of Nanking (now Nanjing) by the Japanese in 1937 (Bergamini, 1971; Dadrian, 2004; Power, 2003; Weitz, 2003). Therefore, Fiske and colleagues’ arguments suggest that the dominant basic emotions will be anger and disgust, associated with the secondary emotions of hate, contempt and resentment.

The bigotry found in atrocities, such as the Holocaust, has been explicitly addressed by Sternberg’s duplex theory of hate (Sternberg, 2003, 2005). This proposes a tripartite structure for hate, comprising passion (for which the concomitant emotions are anger and/or fear), negation of intimacy (for which the concomitant emotion is disgust) and
commitment (which involves cognitions which devalue and diminish the hated object). The risk of atrocity may increase when more types of hate (i.e. different potential combinations of the three components) are present. Hate originates in stories which relate hater and hated, for example by describing the latter as poisoning or seducing the former; these stories increase levels of passion, negation of intimacy or commitment in those who thereby come to hate. In terms of emotion profiles, one would therefore presumably expect that disgust, anger and/or fear predominate in the language of bigotry.

Sternberg emphasizes both negation of intimacy and passion as part of the emotional repertoire of hate, and, by extension, hate crimes. However, this extension is not uncontroversial. Reading accounts of perpetrators, one is often struck by the lack of passion in their descriptions of the people they are killing (Baumeister, 2001; Klee, Dressen, & Riess, 1991). Even in genocide, ‘it is far from clear that genocidal hatred is the usual or primary cause’ (Moshman, 2005, p. 185). This leads to the alternative hypothesis that negation of intimacy is what predominates – implying that disgust may particularly motivate the bigotry associated with atrocities. That is, perpetrators may see the out-groups they target as disgusting social contaminants, and their own behaviour as ‘cleansing’, ‘curing’ or ‘purifying’ the social system.

The role of disgust

This ‘social contamination’ hypothesis suggests that the targeted group may be seen not primarily as a threat to physical survival or to resources, but as a bearer of pollution or disease, a danger to the integrity and purity of an individual or group. Those contaminated themselves become contagious, making them disgusting, dangerous and socially unacceptable. The consequent loss of group support, and the risk of hostility, aggression or expulsion by other members, makes social contamination threats extremely serious. In addition, strangers may bring physical hazards with them, as happened when the populations of what are now Namibia and Congo dropped precipitously during (respectively) German and Belgian colonization (Bridgman & Worley, 2004; Hochschild, 1999).

Disgust may play an important role in generating and maintaining antipathy to out-groups. It is thought to have evolved as a biological rejection mechanism – whether to protect omnivorous human bodies from bad-tasting, potentially dangerous foods (Haidt, McCauley, & Rozin, 1994; Rozin & Fallon, 1987; Rozin, Haidt, & McCauley, 2000) or as a defence against pathogenic disease (Fessler & Narravete, 2005); but its role has expanded to include reactions to moral offences or violations of accepted social behaviour (Haidt, Rozin, McCauley, & Imada, 1997). Out-groups are often accused of such offences, and the language used to describe out-group members often emphasizes disgust. Nazi anti-Semitic propaganda, for example, uses vivid metaphors of disease, contamination and corruption (Gilbert, 1986; Lifton, 2000; Weitz, 2003). Jewishness is described as an infection or cancer, and its ‘carriers’ are likened to rats, maggots, lice and bacilli.

The emphasis on impurity and associated metaphors of hygiene and surgery (Lifton, 2000) are not specific to the Nazis (Dadrian, 2004; Lemarchand, 2002). Their specific symbolic form varies with culture, from the biomedical metaphors of the Nazis to the symbolism of healthy ‘flow’ and unhealthy ‘blockage’ used in Rwanda; but a common theme is an ‘obsessive focus on the body’ (Linke, 2002; Scheper, 2002, p. 368; Taylor, 2002). The concomitant emotion for such an obsession is disgust (Haidt et al.,
Disgust evoked by a threat of contamination prompts action designed to ‘have the offensive thing disappear by the removal of either ourselves or it’ (Miller, 1997, p. 34). The removal of the out-group may therefore become a priority for the in-group when levels of disgust are high (Abrams, 2002).

The social contamination hypothesis predicts increased expressions of disgust in the language used by perpetrators of atrocities and their associates to describe the out-groups they target. However, no increase in anger in anti-group texts is predicted. Anger serves as a threat signal to targets who can read the signal and adjust their behaviour accordingly; dehumanization, however, downplays agency (Cortes, Demoulin, Rodriguez, Rodriguez, & Leyens, 2005) and disgust expects no agentic capability from its targets (Harris & Fiske, 2006). Getting angry with rotting meat would be pointless, whereas disgust is protective, encouraging the person to distance him- or herself from the object’s intrusive unpleasantness (Kolnai, trans. 2004). Human beings labelled as disgusting are seen less as social agents and more as natural hazards, like infections, to be avoided.

Hypotheses
What has been said gives rise to five hypotheses.

Hypothesis 1
The written language used by perpetrators of so-called hate crimes and their ideological sympathizers (authors of anti-group texts) should include words representing multiple negative emotions. In addition, the language of anti-group texts should emphasize some emotions over others (e.g. sadness is not expected to be emphasized). That is, ratings of word frequency in anti-group texts, controlling for word frequency within the language as a whole (assessed using large text corpora as a proxy; see below), should differ significantly between at least some of the basic emotions of anger, sadness, disgust, fear and surprise.

Hypothesis 2
Emotion profiles of anti-group texts will differ from those in the language as a whole.

Hypothesis 3
Emotion profiles in anti-group texts may reflect the perception of the targeted group as a threat to physical survival or resources. In the former case, fear would dominate the profiles; in the latter case, anger, and perhaps also envy and mistrust, would predominate.

Hypothesis 4
The emotion profiles in anti-group texts may depend on the perceived status and competence of the targeted group (Harris & Fiske, 2006). This hypothesis predicts systematic increases in anger, hate and disgust in anti-group texts relative to general language.
Hypothesis 5
The targeted group may be seen primarily as a bearer of contamination. Thus, an increased emphasis on disgust and perhaps hate, but not anger, fear or envy, should be found more in anti-group texts than in general language.

In addition, given the importance of ideological and symbolic factors in both the moralization of disgust (Rozin et al., 2000) and in many atrocities (Lifton, 2000; Taylor, 2002), it may be the case that anti-group texts’ use of metaphors (a term used here, for convenience, to include similes and other analogical language) emphasizes contamination threats over other potential hazards. Words such as cancer and to a lesser extent poison, which are associated with disease and bodily dysfunction, may evoke disgust (Lifton, 2000) and hence should be found more in anti-group texts than in generic language. Metaphors such as swarm and flood, which represent threats to survival but not specifically to the body’s internal integrity, may evoke fear (Canetti, 1973), but are less likely to evoke disgust, and hence serve as a useful comparison. This prediction may distinguish hypotheses 5 and 4, since explanations in terms of status and competence do not necessarily predict the usage of disease-related but not other threat-related metaphors in anti-group texts.

Methods
This study tested these five hypotheses by analysing the prevalence of emotive language in anti-group texts and comparing it with the language as a whole, as assessed using large corpora. For each of the 12 emotion-related words and 12 frequency-matched control words, a set of synonyms was selected from an on-line thesaurus (Roget’s Thesaurus, 2005) and matched against each text in turn. Synonyms were used because the numbers of emotion names themselves in the selected texts were too small to allow statistical analysis. The frequencies of synonyms within the anti-group texts were compared with their frequencies in the language as a whole.

Statistics
Standard statistical software (SPSS, 2001) was used, with statistical significance evaluated at $p < .05$ (two-tailed). Data were largely non-normal (as assessed by Shapiro-Wilk statistics) hence non-parametric tests were used: Mann-Whitney tests when data could reasonably be assumed to be independent and Wilcoxon-signed ranks tests when this assumption was not appropriate (e.g. between texts, as authors of some anti-group texts are known to have influenced others).

Words
Two categories of negative E-words were used. The first comprised E-words chosen because they function as both nouns and verbs (E, $N = 8$; anger, despair (rather than sadness), disgust, envy, fear, hate, mistrust, surprise). Five of these (anger, despair, disgust, fear, surprise) represent basic emotions (Ekman, 1999); envy, mistrust and hate were included because of their clear relevance to prejudice. The second category comprised metaphor words used to denigrate target groups (M, $N = 4$; of which two were postulated to be related to social contamination threats (cancer, poison) and two to more generic threats (flood, swarm)). For both the categories, each E-word was paired with a control word matched on length and frequency, giving a total of 24 words.
Word frequency

The word frequencies for the 24 words used here (see Table 1) were taken from the Kucera and Francis (KF) written corpus (Kucera & Francis, 1967). This is somewhat out-of-date and of limited scope (as it relates to American written texts, whereas some of the texts used in this study are of non-US origin), but it remains the best-known word frequency dataset. For comparison, frequency counts were obtained from the newer and larger British National Corpus (2004a), which includes written and spoken subcorpora, using the SARA programme (British National Corpus, 2004b). The British National Corpus (BNC) and KF measures of word frequency were highly correlated ($r = .92, p < .001$), suggesting that they are equivalent for the purposes of this study.

<table>
<thead>
<tr>
<th>Table 1. Emotion and control words, paired</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotion words</td>
</tr>
<tr>
<td>Pair</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>11</td>
</tr>
<tr>
<td>12</td>
</tr>
</tbody>
</table>

Note. The table shows the emotion and matched control words used in the study. Words are presented in alphabetical order, by type (E, E-word; M, metaphor word). Len, word length; KF, word frequency in the dataset of Kucera and Francis (1967).

Synonyms

Only synonyms from entries which listed the word itself as a main entry were used. Synonym sets included the word itself; they also included multiple-word synonyms, although only single words could be matched against texts. Some synonyms occurred under multiple main entries (e.g. acrimony, which was listed under both bate and anger). Mann-Whitney tests showed, as expected, that the amount of overlap was much greater in the E-words, which come from a common semantic domain, than in the control words ($p < .001$). Emotion and control words did not differ significantly on length, frequency or the number of total, main-entry or multiple-word synonyms.

Generic language: Word frequency measures

The BNC dataset was matched against the list of word synonyms to generate, for each word, a total number of synonym occurrences in the dataset. The frequencies per million for each synonym provided by the BNC (Kilgarriff, 2006) were then summed.
across synonyms to give an overall frequency for each word ($F_{BNC,\text{WORD}}$). For each word in each text, the synonym frequencies per million ($F_{\text{TEXT,WORD}}$) were calculated and compared with the BNC frequency.

**Preprocessing**

In order to compare between words, it was necessary to take account of the varying numbers of word synonyms in the thesaurus ($T_{\text{WORD}}$) and the varying word lengths ($L_{\text{TEXT}}$) of the texts. For each text, the number of synonyms ($S_{\text{WORD}}$) of each word which were present in a text of length $L$ unique words was counted and the normalized variables computed as follows:

$$N_{\text{WORD}} = \left( \frac{S_{\text{WORD}}}{T_{\text{WORD}}} \right) \times 100$$

The $N_{\text{WORD}}$ variables were non-normally distributed for all texts (Shapiro–Wilk statistic in the range $0.59$–$0.87$, $p <= .005$ for all words). They were highly correlated with their corresponding raw synonym counts ($r = .80$, $p < .001$), suggesting that results derived from the $N_{\text{WORD}}$ variables can be generalized to the original data.

**Semantic independence**

Words cannot be assumed to be semantically independent, hence multicollinearity between E-words’ $N_{\text{WORD}}$ ratings may be confounding. If, however, words for basic emotions reflect distinct psychological constructs, as many researchers argue (Cottrell & Neuberg, 2005; Ekman, 1999; c.f. Ortony & Turner, 1990), there should be few significant correlations between E-words in this sample, and those which do exist should be between basic and secondary emotions rather than between one basic emotion and another (e.g. disgust and envy; but not disgust and anger). To address this issue, Spearman correlation coefficients were calculated for $N_{\text{WORD}}$ ratings across all texts. Out of 132 possible E-word pairs, 10 (7.6%) showed statistically significant correlations. None involved two basic emotions, suggesting that multicollinearity is not a significant confound.

**Anti-group texts**

The seven anti-group texts are taken from Internet sources (URLs are given in the references) and are either originally in or translated into English. They represent a variety of political positions, cultures and historical periods, and were selected based on source rather than content by a rater blind to the specific hypotheses set out in this study. All were written either by instigators or perpetrators of violent crimes (e.g. the excerpt from Mein Kampf) or their ideological sympathizers, or by authors whose expressed opinions are politically ‘extremist’ in that adherents of such opinions have used violence to pursue their political goals.

In particular, these texts express extreme prejudice and often advocate violence against one or more target groups. This is not, of course, to allege that the authors themselves advocate or participate in violence, although some (e.g. Adolf Hitler) have done so. Targeted groups are: Muslims and their allies in British society (Text 1), the West (Text 2, Text 7), ‘non-white races’ (Text 3), women (Text 4) and Jews (Text 3, Text 5, Text 6). The texts were also chosen to sample varying degrees of emotional tone: some, for example Text 3, read as relatively matter-of-fact; others (e.g. Text 2) are
intensely emotive (this variability is reflected in rater responses; see Table 2). The groups involved also vary in their political influence, from the relatively low levels achieved by the British National Party (Text 1) through to the domination of an entire society, and consequent genocide, in Nazi Germany (Text 5).

Table 2. Text assessment questionnaire items

<table>
<thead>
<tr>
<th>Text assessment questionnaire items</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Would you describe the author of the text as an effective politician?</td>
<td>3.29</td>
<td>0.60</td>
</tr>
<tr>
<td>2 Would you say the text expresses negative prejudice(s)?</td>
<td>4.36</td>
<td>0.10</td>
</tr>
<tr>
<td>3 Would you say the text expresses mainstream political opinions?</td>
<td>2.00</td>
<td>0.20</td>
</tr>
<tr>
<td>4 Would you describe the opinions expressed in the text as bigoted?</td>
<td>4.22</td>
<td>0.50</td>
</tr>
<tr>
<td>5 Assuming the text reflects its author’s opinions, would you vote for the author if he/she were to stand for Parliament in a general election?</td>
<td>1.15</td>
<td>0.21</td>
</tr>
<tr>
<td>6 While reading, did you feel that the author was attempting to manipulate your emotions?</td>
<td>4.14</td>
<td>0.81</td>
</tr>
<tr>
<td>7 How emotive did you find the text?</td>
<td>3.15</td>
<td>1.62</td>
</tr>
<tr>
<td>8 To what extent you think the text is about a certain group of people (however defined), or about the culture, behaviour or ideas associated with them?</td>
<td>4.50</td>
<td>0.51</td>
</tr>
<tr>
<td>9 In the text, how would you describe the attitude to (insert name of target group)?</td>
<td>1.22</td>
<td>0.11</td>
</tr>
<tr>
<td>10 Does the text advocate, condone or encourage violence against (insert name of target group)?</td>
<td>3.71</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Note. The table shows the 10 questionnaire items used by two independent raters to assess the anti-group texts. Means and standard deviations (SD) of rater responses are shown in the right-hand columns.

Text ratings

Two independent raters read the texts and assessed them, using 5-point Likert scales, on a 10-item questionnaire. The items included questions about the texts (e.g. the extent to which they appeared to express prejudice, bigotry, hostility to the target group and pro-violence attitudes) and their authors, who were not identified (e.g. political ability, position relative to mainstream politics). Items are listed in Table 2, together with mean ratings and standard deviations. One item, relating to pro-violence attitudes, had 0 variance; for the other nine items, Cronbach’s $\alpha = .76$.

Raters were also asked to generate up to five terms to describe the emotions which they felt each text expressed and up to five more listing the emotions they felt the author was trying to arouse in them. These words were classified into the following categories: hatred (e.g. ‘anti-Semitic feeling’), anger (e.g. ‘fury’), disgust, fear and anxiety, sadness, shame and guilt, other negative emotions, positive emotions, and a category including such terms as ‘sternness’, ‘rigour’, ‘fanaticism’, ‘fascism’, ‘dogmatism’, ‘fundamentalism’ and ‘willingness to take violent action’, which appears to reflect strength of belief/commitment.

In addition, 51 out of the 84 words listed by the raters, or close cognates (e.g. ‘disgusted’), have been assessed on 9-point scales (positive-negative) of valence and arousal as part of the Affective Norms for English Words (ANEW) project (Bradley &
Lang, 1999). Where available, these measures were recorded and means/standard deviations calculated.

Finally, for each occurrence of an E-word in a text, the semantic context was assessed as congruent or incongruent with the valence of the emotion of which the E-word was a synonym.

Text descriptions

Text descriptions are as follows:

- **Text 1.** British National Party web site, article (Barnes, 2004). The BNP is an extreme nationalist party on the fringe of British politics, several of whose leaders have been convicted of crimes including incitement to racial hatred (Wikipedia, 2005). The text in question, which was written before the London bombings on 7 July 2005, targets 'Islamic terrorists' and apologists for the Islamic community', who are criticized for spreading the belief that terrorism is perpetrated by the socio-economically disadvantaged.

- **Text 2.** Islam On-line web site, article (Islam On-line, 2004). The target is rarely named, but is clearly the West/Israel and particularly the current US administration. Atrocities and humiliations such as Abu Ghraib are discussed. The article’s political position is close to that of many Islamist radical thinkers who advocate violence against Western targets.

- **Text 3.** Kingdom Identity web site doctrinal statement (Kingdom Identity, 2005). The target of this self-described ‘politically incorrect’ Christian Identity sect is ‘non-white races’ and the Jews, whose moral inferiority and descent from a corrupt ‘seedline’ is contrasted with the spiritual and genetic purity of the chosen White race. No author is listed.

- **Text 4.** *Malleus Maleficarum* (‘The Hammer of Evildoers’), extracts from pp. 11–125 (Malleus Maleficarum, trans. 1928). This hugely influential book was written in 15th century Europe as a guide to the identification and treatment of witches (mainly women). Estimates of the number of people killed during the European witch-crazes vary widely but may be around 50,000 (Briggs, 1996). The extracts describe women’s moral weaknesses, treachery, emotionality and tendency to sin (e.g. by copulating with devils). The sections less relevant to the overall argument (e.g. Biblical examples) have been omitted.

- **Text 5.** *Mein Kampf*, Chapter 11 (Hitler, trans. 1939). This is the notorious chapter on race and people, in which Hitler sets out his view of how ‘the Jew’ has operated as a social parasite throughout history.

- **Text 6.** *The Mystery of Iniquity*, by Wesley Swift (Swift, 2005). The target of this sermon by a key figure in the Christian Identity movement is the ‘mystery of iniquity’, described by St. Paul in 2 Thessalonians and identified by Swift as the Jews. Swift gives a historical description of alleged Jewish atrocities and undermining of Christianity. The date of the text is not given, though it is prior to 1970, when Swift died.

- **Text 7.** Writings of Sayyid Qutb: Chapter 10 of his book Ma’alim fi l-Tariq, first published in 1964 and rendered into English as *Milestones* (Qutb, trans. 2003). Qutb is a foundational figure in modern Islamic radicalism (Ruthven, 2004). His target is ‘jahiliyyah’, or ignorance, as exemplified in Qutb’s view by the West, particularly its consumerist, individualist culture. Qutb claimed that Muslim governments had also fallen into jahiliyyah and that their authority was therefore illegitimate. He was executed by the Egyptian regime in 1966.
Preprocessing
Each text was processed, using a combination of automated techniques and visual inspection, to place each word on a single line and remove punctuation, paragraph spacing, numbers, quotation marks, indents and apostrophes (including ‘s). Biblical and other source references, usually given in brackets, were also removed prior to text analysis. Titles, names and place names comprising multiple words (e.g. Adolf Hitler) were treated as single word units.

Analysis
To assess whether anti-group texts contained negative emotional language, whether word frequencies in anti-group texts differed between emotions and whether any effects were specific to the E-words, N\textsubscript{WORD} ratings for the eight E-words (anger, despair, disgust, envy, fear, hate, mistrust, surprise) and the eight frequency-matched control words were compared pairwise using Wilcoxon tests. Visual inspection was used to assess the direction of any significant contrast. To compare the frequencies of E-words (including synonyms) in anti-group texts (F\textsubscript{TEXT,WORD}) and in the language as a whole (F\textsubscript{BNC,WORD}), Mann–Whitney tests were used (see Table 3).

Results
Ratings of anti-group texts
Raters were asked to assess, on a scale of 1 to 5, various aspects of the anti-group texts’ content (see Table 2 for details). Text were rated as variably emotive (mean = 3.15), negatively prejudiced (mean = 4.26), bigoted (mean = 4.22) and emotionally manipulative (mean = 4.14). They were seen as targeting a certain group or groups (mean = 4.50), expressing negative attitudes to those groups (mean = 1.22; 1 = ‘extremely negative’, 5 = ‘extremely positive’), and as advocating, condoning or encouraging violence against them (mean = 3.71). Raters were also asked about political aspects of the texts. They tended to see the authors as effective politicians (mean = 3.29), but not as politically mainstream (mean = 2.00) and not as deserving their vote (mean = 1.15).

Out of the 84 (41 unique) emotion terms produced by the raters, ANEW ratings were available for 51. These indicated negative valence (mean = 2.68, standard deviation (SD) = 1.75) and high arousal (mean = 6.78, SD = 1.00). Out of the 84 terms, 54 (64%) were either ‘hatred’ (13), ‘anger’ (10), ‘disgust’ (9) or related emotions (e.g. ‘misogyny’, ‘sense of injustice’, ‘disapproval’, ‘frustration’, ‘humiliation’). Horror was mentioned once, while only four terms (5%) referred to fear or anxiety and none referred to other negative emotions, such as shame, guilt or sadness. Six terms (7%) were positive. The remaining 19 terms (23%) were related to strength of belief, conviction and dogmatism. The predominance of negatively valenced words is thus restricted to particular emotions.

The valence-congruence or otherwise of each E-word occurrence in the texts was also assessed by considering its semantic context. E-words for disgust, hate, cancer and poison were valence-congruent in all cases. The E-words for anger were valence-congruent with the following possible exceptions: one occurrence of the synonym ‘excite’ (the phrase, ‘to excite her husband with hot words’, comes in a list of immoral female behaviour) and five references to Christ’s Cross or Passion. The E-words for the other seven words are more often used valence-incongruently (e.g. ‘wonder’ as a synonym for mistrust).
Table 3. Statistics for emotion words, by text group.

<table>
<thead>
<tr>
<th>Word</th>
<th>Anger</th>
<th>Cancer</th>
<th>Despair</th>
<th>Disgust</th>
<th>Envy</th>
<th>Fear</th>
<th>Flood</th>
<th>Hate</th>
<th>Mistrust</th>
<th>Poison</th>
<th>Surprise</th>
<th>Swarm</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Anti-group text word frequencies (per million))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>1,618</td>
<td>1,469</td>
<td>690</td>
<td>1,013</td>
<td>2,234</td>
<td>871</td>
<td>439</td>
<td>2,069</td>
<td>701</td>
<td>1,212</td>
<td>1,654</td>
<td>904</td>
</tr>
<tr>
<td>SE</td>
<td>515</td>
<td>326</td>
<td>193</td>
<td>183</td>
<td>649</td>
<td>335</td>
<td>239</td>
<td>573</td>
<td>213</td>
<td>306</td>
<td>260</td>
<td>355</td>
</tr>
<tr>
<td>SD</td>
<td>1,364</td>
<td>863</td>
<td>511</td>
<td>485</td>
<td>1,718</td>
<td>887</td>
<td>633</td>
<td>1,517</td>
<td>563</td>
<td>811</td>
<td>688</td>
<td>938</td>
</tr>
<tr>
<td>BNC word frequencies (per million)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1,267</td>
<td>446</td>
<td>988</td>
<td>494</td>
<td>2,970</td>
<td>3,225</td>
<td>1,503</td>
<td>1,035</td>
<td>2,352</td>
<td>938</td>
<td>5,736</td>
<td>2,890</td>
</tr>
<tr>
<td>Change in anti-group texts relative to BNC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ρ</td>
<td>↑</td>
<td>↑</td>
<td>↓</td>
<td>↑</td>
<td>↓</td>
<td>↓</td>
<td>↑</td>
<td>↓</td>
<td>↑</td>
<td>↓</td>
<td>↓</td>
<td>↑</td>
</tr>
</tbody>
</table>
| Note: For the 12 emotion words used in the study (listed alphabetically), the table shows the mean, standard error of the mean (SE), and standard deviation (SD) of the frequencies per million for the anti-group texts (upper panel). The middle panel shows frequencies for the BNC corpus. The lower panel gives the Mann-Whitney U statistic (MW) and the p value (p) for a comparison of the two text groups. Upward arrows indicate higher levels of a word and its synonyms in the anti-group texts than in the BNC corpus, downward arrows indicate lower levels. ns, non-significant (p > .05). *, significant at the level p < .05; **, significant at the level p < .01.
Taken together, these findings imply an emphasis within the anti-group texts on the emotions of anger, disgust and hatred. Other negative emotions are not emphasized; where their synonyms occur, they are less likely to be used with a valence-congruent meaning.

**Emotion profiles in anti-group texts**

$N_{\text{WORD}}$ ratings for the eight E-words (*anger*, *despair*, *disgust*, *envy*, *fear*, *hate*, *mistrust*, *surprise*) were compared pairwise using Wilcoxon tests. If emotion profiles are not differentiated within anti-group texts, the results should indicate uniform non-significance, as the $N_{\text{WORD}}$ ratings control for frequency in the language in general. In fact, as predicted by Hypothesis 1, differences were found: for *envy* and *disgust* ($\text{Wilcoxon } Z = -2.03, p = .043$); *fear* and *disgust* ($\text{Wilcoxon } Z = -2.37, p = .018$); *fear* and *envy* ($\text{Wilcoxon } Z = -2.20, p = .028$); *hate* and *fear* ($\text{Wilcoxon } Z = -2.20, p = .028$); *mistrust* and *envy* ($\text{Wilcoxon } Z = -1.99, p = .046$) and *surprise* and *envy* ($\text{Wilcoxon } Z = -2.03, p = .043$).

**E-word counts in anti-group texts compared with general language**

None of the texts is listed as a source for the BNC corpus, hence texts and BNC were regarded as statistically independent. For each emotion and metaphor word, Mann-Whitney tests were therefore used to compare word frequencies per million based on counts in anti-group texts and in the language as a whole, using the BNC as a proxy. Despite the small sample size, the anti-group texts showed significant increases in language related to *disgust* and *hate* ($\text{Wilcoxon } Z = 2.16, p = .03$) and the metaphor word *cancer* ($\text{Wilcoxon } Z = 2.51, p < .01$).

Increases in *anger* and *poison* were non-significant. Other E-words appeared significantly less often in the anti-group texts than in the BNC, as did all the control words with the exception of *delusion*, *prior* and *reject*, where differences were non-significant. The lower levels may reflect the more specialized content of the anti-group texts. The non-significant result for *poison* may, speculatively, be related to its lesser efficacy as a disgust metaphor; it lacks *cancer*’s connotations of squamous, uncontrollable organicity (Miller, 1997).

Figure 1 shows the mean percentage difference in frequency between the anti-group texts and the BNC corpus. Values greater than 0, as shown by bar height, indicate increases; values below 0 indicate decreases. Significant increases are marked by an asterisk above the bar.

**Hypothesis 1**

Hypothesis 1 states that in anti-group texts E-words should show a non-uniform emotion profile, controlling for their frequency in the language as a whole. The results suggest that there are indeed significant differences, for example between *fear* and *disgust*. Hypothesis 1 is therefore supported.

**Hypothesis 2**

Hypothesis 2 states that anti-group texts will have distinct emotion profiles from language in general. Results suggest that systematic differences in emotional language do distinguish the more extreme anti-group texts from generic language. Hypothesis 2 is therefore supported.
Hypotheses 3–5

Hypothesis 3 states that anti-group texts differ from generic language in having higher prevalences of fear or anger, and possibly envy and mistrust. Hypothesis 4 states that anti-group texts will have higher prevalences of disgust, anger and hate. Hypothesis 5 states that anti-group texts will have higher prevalences of disgust and perhaps hate; additionally, cancer and possibly poison, but not swarm and flood, will be more prevalent in anti-group texts.

The results do not provide much support for Hypothesis 3; there is little change in anger, and fear-related words are less prevalent than in generic language. Hypothesis 4 is supported in that there is some evidence for increases in hate and disgust, although not for anger. Hypothesis 5 also receives support from this finding. In addition, the pattern of changes for the metaphor words is as predicted by Hypothesis 5.

Discussion

This study shows that the synonyms of words (including metaphors) related to disgust and hate are more prevalent in anti-group texts than in general language. These findings imply that the evocation of high levels of disgust may be a crucial motivational factor in bigotry and perhaps in the perpetration of intergroup atrocities. Thus, they support the social contamination hypothesis. They are also consistent with the model proposed by Fiske and colleagues, although this does not
specifically predict the differences in metaphor words. The findings do not support hypotheses which emphasize the role of anger and fear in atrocities. In terms of Sternberg’s duplex theory of hate, negation of intimacy, rather than passion, appears to dominate the discourses assessed here.

This study makes no comment on the cognitive aspects of atrocities, other than to note the prevalence of terms relating to strength of belief in raters’ descriptions of the anti-group texts (again, this brings to mind Sternberg’s construct of commitment). Adherence to false and damaging beliefs appears to be high in at least some perpetrators, though perhaps not all (Browning, 1991). It is worth noting, however, that the use of the metaphor word cancer and its synonyms was significantly higher in anti-group texts. This is consistent with Sternberg’s emphasis on the denigratory stories told about target groups.

**Contribution of this research**
The findings described here clearly confirm the rich affective variety within the traditional concept of prejudice. Additionally, this study makes a twofold contribution to the research literature. First, the methodology moves beyond the necessary but artificial constraints of the laboratory to allow the assessment of real-world anti-group texts. Social desirability considerations are reduced, as the texts were written primarily to be read by ideologically sympathetic members of the authors’ in-groups. Secondly, this study differs from most research on prejudice in addressing the negative extreme of intergroup hostility.

**Limitations**
Clearly, the methodology used here has limitations. For one thing, the results may depend on the texts selected. The aim of this study was to compare a range of texts in the hope that including a variety of styles and genres, as well as political positions and historical periods, would make the results more generalizable. Any such choice contains a subjective judgment of what is extreme; moreover, the sample is small. Nevertheless, all the texts classified as ‘anti-group’ in this study clearly promote intergroup hostility, and all the groups whose ideologies they represent have been associated with ideologically motivated intergroup violence, from assaults and murders of individuals, through terrorist atrocities such as 9/11, to state-mediated mass killings like the medieval witch-hunts and the Holocaust.

Another potential criticism of this study is that it has not tested the assumption that synonym use reflects ideological intent, perhaps as well as or better than the use of the word itself. Access to the intentions of the authors of the texts used here was not available. However, if there were no such relationship between synonym use and intention, there would be no reason to suppose that any difference should affect E-words unevenly. Instead, the finding of significant differences for disgust and hate in anti-group texts is arguably strengthened by the fact that the finding is specific to those emotions. Other emotions and frequency-matched control words are not significantly more prevalent in anti-group texts.

No attempt has been made to take account of variations in the cultural origin of the texts. Nor does the study consider variations in the power relations between the in-groups whose members wrote the anti-group texts and the out-groups.
those texts attack. Further research on larger samples is required to address these
issues.

Finally, and in line with most research on prejudice, only negative emotions have been
considered. Nevertheless, unpalatable though it may be to consider the joys of bigotry
(Billig, 2002a; Frosh, 2002), such joys are not restricted to sadists and psychopaths. They
also occur in non-clinical groups (Browning, 1991; Dutton et al., 2005; Haney, Banks, &
Zimbardo, 1983) – if cruelty is committed in a group context (Nell, 2006).

**The social contamination hypothesis**

If disgust is indeed involved in atrocities, then the language of hygiene used by many
perpetrators is no coincidence. Rather, they may actually see themselves as the immune
system of a wider ‘body politic’ which must be defended from dangerous social
pathogens. Individuals for whom the avoidance of disgusting stimuli is particularly
salient, either because of personal characteristics or because their profession
emphasizes such avoidance (e.g. health professionals), may be particularly likely to
find such metaphors appealing (Lifton, 2000). These may be the political entrepreneurs
who help facilitate violence (Tilly, 2003), in part by encouraging others to think in terms
of cleanliness and pollution rather than, say, threat and response, or social or economic
competitors.

If the processes which lead to atrocities are understood in terms of social
contamination, then the following five predictions can be made:

**Prediction 1:** in-group membership should become more salient and in-group–out-
group polarization more extreme when individuals have been primed to feel disgust
than when they have been primed to feel anger.

**Prediction 2:** minority groups who transgress (or who are believed to transgress)
majority social codes in disgust-sensitive domains – for example food, sex, waste disposal
– are more likely to be targets of extreme hostility, atrocities and genocide than those who
do not. Propaganda focusing on these areas will be a feature of the build-up to atrocities.

**Prediction 3:** perpetrator training prior to planned atrocities should include what
one may term ‘disgust ordeals’, deliberately confronting the trainee with disgusting
stimuli while providing highly rewarding group feedback.

**Prediction 4:** perpetrator behaviour during atrocities should include attempts to
ensure that the victims are seen as disgusting, especially when perpetrator behaviour is
challenged (e.g. by particularly vulnerable and pathetic victims, or perpetrator
inexperience). This may help to explain some of the worst excesses observed during
atrocities, such as the example cited by Holocaust scholar Lawrence Langer of an SS
man’s behaviour during a *KinderAktion* – a roundup of young people for execution – in
the Jewish ghetto of Kovno (now Kaunas) in Lithuania. A survivor of the ghetto testified
to being ‘present in the room when an SS man entered and demanded from a mother the
one-year-old infant she was holding in her arms. She refused to surrender it, so he seized
the baby by its ankles and tore the body in two before the mother’s eyes’ (Langer, 1999,
p. 3). Such behaviour is normally interpreted as gratuitous savagery – the SS man could,
after all, simply have broken the baby’s neck. However, the social contamination
hypothesis suggests an alternative explanation. By tearing the child apart, the SS man
transformed it from a human infant not into a dead baby – still clearly human and thus
deserving of grief and sympathy – but into a dismembered corpse, evoking revulsion,
thereby reinforcing the ‘correct’ Nazi attitude to victims.
Prediction 5: when it comes to the thorny issue of atrocity prevention, strategies which focus on reducing disgust (e.g. by having respected leaders openly engage in physical contact with members of the targeted group) may be more effective than those which aim to diminish anger or fear.

Concluding remarks

Clearly, the findings described in this study are only preliminary. Much more research is required using other texts and methodologies, and numerous issues remain to be addressed. Nevertheless, the specificity of the findings is provocative. When considering the difficult topic of intergroup atrocities, perhaps the most difficult problem in social psychology, the hitherto relatively neglected emotion of disgust may reward investigation.

Acknowledgements

Many thanks to Professors John Dovidio and Rupert Brown, and to two unnamed reviewers, for their critical feedback.

References


In D. M. Mackie & E. R. Smith (Eds.), From prejudice to intergroup emotions: Differentiated reactions to social groups (pp. 89–109). New York: Psychology Press.


Disgust and extreme prejudice 615


From prejudice to intergroup emotions: Differentiated reactions to social groups. New York: Psychology Press.


Received 22 February 2006; revised version received 25 August 2006.